DUKE MATHEMATICAL JOURNAL

DUKE UNIV. LIBRARY
DURHAM, N. C.

DUKE

EDITED BY

LEONARD CARLITZ DAVID VERNON WIDDER

JOSEPH MILLER THOMAS

Managing Editor

WITH THE COÖPERATION OF

R. P. BOAS, JR.	NATHAN JACOBSON	A. J. MARIA	J. A. SHOHAT
H. E. BRAY	C. G. LATIMER	F. J. MURRAY	R. J. WALKER
J. J. GERGEN	NORMAN LEVINSON	S. B. MYERS	MORGAN WARD
G. A. HEDLUND	E. J. McSHANE	OYSTEIN ORE	G. T. WHYBURN
	C. C. MacDUFFEE	J. H. ROBERTS	

Volume 5, Number 1

MARCH, 1939

COPYRIGHT, 1939

DUKE UNIVERSITY PRESS

DURHAM, N. C.

DUKE MATHEMATICAL JOURNAL

This periodical is published quarterly under the auspices of Duke University by Duke University Press at Durham, North Carolina. It is printed at Mt. Royal and Guilford Avenues, Baltimore, Maryland, by the Waverly Press.

The American Mathematical Society is officially represented on the Editorial Board by Professors Ore and Whyburn.

The subscription price for the current volume is four dollars; single numbers, one dollar and twenty-five cents; back volumes, five dollars each. Individual members of the Mathematical Association of America may subscribe at half price. Mention of such membership should be made when subscribing. Subscriptions, orders for back numbers, and notice of change of address should be sent to Duke University Press, Durham, North Carolina.

Manuscripts submitted for consideration by the editors should be typewritten, with the exception of the formulas, which should be filled in by hand. It is recommended, though not required, that footnotes be numbered consecutively with arabic numerals, and that they be placed at the end of the manuscript. A footnote will appear, however, at the bottom of the printed page of the journal on which reference is made to it. The author should keep a duplicate copy of the manuscript as a precaution against loss in transit and as an aid in reading proof. The original manuscript will not accompany the galley proof sent to the author unless requested. Such a request may delay the appearance of the article.

Authors are entitled to one hundred free reprints. Additional copies will be furnished at cost. All reprints will be furnished with covers unless the contrary is specifically requested.

Manuscripts and editorial correspondence should be addressed to Duke Mathematical Journal, 4785 Duke Station, Durham, North Carolina.

Entered as second class matter at the Post Office, Durham, North Carolina. Additional entry at the Post Office, Baltimore, Maryland.

Made in United States of America

WAVERLY PRESS, INC.
BALTIMORE, U. S. A.

THE ERGODIC THEOREM

By Norbert Wiener

1. Ergodic theory has its roots in statistical mechanics. Both in the older Maxwell theory and in the later theory of Gibbs, it is necessary to make some sort of logical transition between the average behavior of all dynamical systems of a given family or ensemble, and the historical average behavior of a single system. This transition was not carried out with any rigor until the theory of Lebesgue measure had been developed. The fundamental theorems are those due to Koopman, von Neumann, and Carleman, on the one hand, and to Birkhoff, on the other. A careful account of them and their proofs is to be found in Eberhard Hopf's *Ergodentheorie* (Berlin, 1937) in the series *Ergebnisse der Mathematik und ihrer Grenzgebiete*. The bibliography of that monograph is so complete that it relieves me from all need of furnishing one on my own account. It is important to point out, however, that Birkhoff's first paper (Proc. Nat. Acad. Sci. (1931)) contains Theorem IV of this paper.

The fundamental theorems of ergodic theory, which emerged in the epoch 1931–32, are recognized as theorems pertaining in the first instance to the abstract theory of the Lebesgue integral. Much of this paper will be devoted to new proofs of these known results and to their proper orientation mutually and with respect to the rest of analysis. They have several variant forms, but perhaps the simplest form of the von Neumann theorem reads:

Theorem I. *Let S be a measurable set of points of finite measure. Let T be a transformation of S into itself, which transforms every measurable subset of S into a set of equal measure, and whose inverse has the same property. Let $f(P)$ be a function defined over S and of Lebesgue class L^2. Then there exists a function $f_1(P)$, also belonging to L^2 and such that*

$$\lim_{N\to\infty} \int_S \left| f_1(P) - \frac{1}{N+1} \sum_{n=0}^{N} f(T^n P) \right|^2 dV_P = 0. \tag{1.01}$$

Birkhoff's theorem reads:

Theorem II. *Let S and T be as in Theorem* I. *Let $f(P)$ be a function defined over S and of Lebesgue class L. Then, except for a set of points P of zero measure,*

$$f_1(P) = \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} f(T^n P) \tag{1.02}$$

will exist and belong to L.

These theorems have continuous analogues. These are, respectively:

Received December 7, 1938; an invited address before the American Mathematical Society at its meeting held April 7-8 in conjunction with the centennial celebration of Duke University.

377070

THEOREM I'. *Let S be as in the hypothesis of Theorem* I. *Let T^λ satisfy the conditions given for T in the hypothesis of Theorem* I, *for every real value of λ, and let*

$$T^\lambda(T^\mu P) = T^{\lambda+\mu}P. \tag{1.03}$$

Let $f(P)$ be a function defined over S and of Lebesgue class L^2, and let $f(T^\lambda P)$ be measurable in the product space of λ and P. Then there exists a function $f_1(P)$, also belonging to L^2 and such that

$$\lim_{N\to\infty} \int_S \left| f_1(P) - \frac{1}{N}\int_0^N f(T^\lambda P)\, d\lambda \right|^2 dV_P = 0. \tag{1.04}$$

THEOREM II'. *Let S and T^λ be as in the hypothesis of Theorem* I'. *Let $f(P)$ be a function defined over S and of Lebesgue class L, and let $f(T^\lambda P)$ be measurable in the product space of λ and P. Then, except for a set of points P of zero measure,*

$$f_1(P) = \lim_{N\to\infty} \frac{1}{N}\int_0^N f(T^\lambda P)\, d\lambda \tag{1.05}$$

will exist as a function in L.

Theorem II' has a very close formal analogy to the fundamental theorem of the calculus. One form of the latter reads:

THEOREM III. *Let $f(x)$ belong to L. Then, for all values of x with the exception of a set of zero measure,*

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon}\int_0^\epsilon f(x+\lambda)\, d\lambda = f(x). \tag{1.06}$$

This analogy becomes even closer if we substitute for Theorem III

THEOREM III'. *On the hypothesis of Theorem* II', *except for a set of points P of zero measure,*

$$f(P) = \lim_{\epsilon\to 0} \frac{1}{\epsilon}\int_0^\epsilon f(T^\lambda P)\, d\lambda. \tag{1.07}$$

It will be seen that the only difference between the pattern of Theorem II' and Theorem III' is that in the first case $N \to \infty$, while in the second case $\epsilon \to 0$. This suggests the existence of a central theorem in which we suppose neither the one nor the other, but deal with $\frac{1}{A}\int_0^A f(T^\lambda P)\, d\lambda$ for all values of A. Birkhoff in fact has proved a theorem of this type.[1] A sharper form of it is the following:

THEOREM IV. *On the hypothesis of Theorem* II, *let $f(P) \geqq 0$ on S and let*

$$f^*(P) = \text{l.u.b.}_{0<A<\infty} \frac{1}{A+1}\sum_{n=0}^{A} f(T^n P); \tag{1.08}$$

[1] Cf. G. D. Birkhoff, Proc. Nat. Acad. Sci., vol. 17(1931), pp. 650-660; also N. Wiener, *The homogeneous chaos*, Amer. Jour. Math., vol. 60(1938), p. 907.

or on the hypothesis of Theorem II′, *let*

$$f^*(P) = \text{l.u.b.}_{0<A<\infty} \frac{1}{A} \int_0^A f(T^\lambda P)\, d\lambda. \tag{1.09}$$

In either case, if $\alpha > 0$, *the measure of the set of points* P *for which* $f^*(P) \geqq \alpha$ *does not exceed*

$$\frac{1}{\alpha} \int_S f(P)\, dV_P. \tag{1.10}$$

It also does not exceed

$$\frac{2}{\alpha} \int_{f(P)\geqq\frac{1}{2}\alpha} f(P)\, dV_P. \tag{1.11}$$

The importance of Theorem IV has been much neglected in the subsequent literature and can be properly appreciated only by a direct reference to Birkhoff's own work. As a corollary of Theorem IV, we have

THEOREM V. *If we define* $f^*(P)$ *as in Theorem* IV, *then if* $f(P)$ *belongs to* L^p $(p > 1)$, *so does* $f^*(P)$; *while if*

$$\int_S f(P) \log^+ f(P) dV_P < \infty, \tag{1.12}$$

then $f^*(P)$ *belongs to* L.

This last theorem is intimately connected with an inequality of Hardy and Littlewood.[2]

All of the theorems so far quoted have analogues in which the group T^λ is replaced by an Abelian group with more than one generator. Theorem I becomes

THEOREM I″. *Let* S *be a measurable set of points of finite measure. Let* $T_1^{\lambda_1} T_2^{\lambda_2} \cdots T_n^{\lambda_n}$ *satisfy the conditions given for* T *in the hypothesis of Theorem* I, *for every set of real values of* $(\lambda_1, \cdots, \lambda_n)$, *and let*

$$T_1^{\lambda_1} T_2^{\lambda_2} \cdots T_n^{\lambda_n} (T_1^{\mu_1} \cdots T_n^{\mu_n} P) = T_1^{\lambda_1+\mu_1} \cdots T_n^{\lambda_n+\mu_n} P. \tag{1.13}$$

Let $f(P)$ *be a function defined over* S *and of Lebesgue class* L^2, *and let* $f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P)$ *be measurable in the product space of* $(\lambda_1, \cdots, \lambda_n)$ *and* P. *Then there exists a function* $f_1(P)$, *also belonging to* L^2 *and such that*[3]

$$\lim_{A\to\infty} \int_S \left| f_1(P) - \frac{1}{V(A)} \int_{\lambda_1^2+\cdots+\lambda_n^2\leqq A^2} \cdots \int f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P)\, d\lambda_1 \cdots d\lambda_n \right|^2 dV_P = 0. \tag{1.14}$$

[2] Cf. Hardy, Littlewood, and Pólya, *Inequalities*, Cambridge, 1934.

[3] Here we use $V(r)$ for the volume of a sphere of radius r in n-space.

377070

Theorem II becomes

THEOREM II''. *Let S and $T_1^{\lambda_1} \cdots T_n^{\lambda_n}$ be as in the hypothesis of Theorem* I'', *and let $f(P)$ be a function defined over S and of Lebesgue class L. Let $f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P)$ be measurable in the product space of $(\lambda_1, \cdots, \lambda_n)$ and P. Then, except for a set of points P of zero measure,*

$$f_1(P) = \lim_{\Delta \to \infty} \frac{1}{V(\Delta)} \int_{\lambda_1^2 + \cdots + \lambda_n^2 \leq \Delta^2} \cdots \int f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P)\, d\lambda_1 \cdots d\lambda_n \tag{1.15}$$

will exist.

Theorem III' will become

THEOREM III''. *On the hypothesis of Theorem* III', *except for a set of points P of zero measure,*

$$f(P) = \lim_{\Delta \to 0} \frac{1}{V(\Delta)} \int_{\lambda_1^2 + \cdots + \lambda_n^2 \leq \Delta^2} \cdots \int f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P)\, d\lambda_1 \cdots d\lambda_n. \tag{1.16}$$

Theorem IV will become

THEOREM IV'. *On the hypothesis of Theorem* II'', *let*

$$f^*(P) = \text{l.u.b.}_{0 < \Delta < \infty} \frac{1}{V(\Delta)} \int_{\lambda_1^2 + \cdots + \lambda_n^2 \leq \Delta^2} \cdots \int f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P)\, d\lambda_1 \cdots d\lambda_n. \tag{1.17}$$

If $\alpha > 0$, the measure of the set of points P for which $f^*(P) \geqq \alpha$ does not exceed

$$\frac{B_n}{\alpha} \int_S f(P)\, dV_P, \tag{1.18}$$

where B_n depends only on the number n of dimensions. It also does not exceed

$$\frac{2B_n}{\alpha} \int_{f(P) \geq \frac{1}{2}\alpha} f(P) dV_P. \tag{1.19}$$

The generalization of Theorem V has verbally the same statement as Theorem V itself, with the exception that the definition of $f^*(P)$ is made as in Theorem IV' instead of as in Theorem IV. We shall call this theorem *Theorem* V'.

The order of proof will be the following: we shall first establish the theorems of the I type.[4] Then we give a completely autonomous proof of the theorems of the IV type. Theorems of the III and the V type result from theorems of the IV type alone, while theorems of the II type result from the combination of theorems of the I and the IV type.

[4] Theorem I has been established in a very direct way in a recent unpublished paper by F. Riesz, and Theorem I'' by Dunford.

2. In the proof of theorems of the I type, the following lemmas are of great service:

LEMMA A. *Let $\phi_k(x)$ belong to L^2 for $1 \leqq k \leqq n$. Let*

$$\int |\phi_k(x)|^2\,dx = A \tag{2.01}$$

for all admissible values of k. Then

$$\int \left|\frac{1}{n}\sum_1^n \phi_k(x)\right|^2 dx \leqq A. \tag{2.02}$$

LEMMA A′. *Let $\phi(x, y)$ be measurable in x and y, and let*

$$\int |\phi(x, y)|^2\,dx = A \qquad \text{for all } y. \tag{2.03}$$

Let $\psi(y)$ belong to L. Then $\int \phi(x, y)\psi(y)\,dy$ belongs to L^2, and

$$\int \left|\int \phi(x, y)\psi(y)\,dy\right|^2 dx \leqq \left[\int |\psi(y)|\,dy\right]^2 A. \tag{2.04}$$

LEMMA B. *Let $\phi_k(x)$ be a set of functions of L^2 for $1 \leqq k \leqq n$. Then there exists a value of k, say k_1, such that*

$$\int \left|\phi_{k_1}(x) - \frac{1}{n}\sum_{k=1}^n \phi_k(x)\right|^2 dx \leqq \int |\phi_{k_1}(x)|^2\,dx - \int \left|\frac{1}{n}\sum_{k=1}^n \phi_k(x)\right|^2 dx. \tag{2.05}$$

LEMMA B′. *Let $\phi(x, y)$ be measurable in x and y, and let $\int |\phi(x, y)|^2\,dx$ be uniformly bounded in y. Let $\psi(y)$ be a non-negative function of class L such that $\int \psi(y)\,dy = 1$. Then there exists a value of y, say y_1, such that*

$$\begin{aligned}\int \left|\phi(x, y_1) - \int \phi(x, y)\psi(y)\,dy\right|^2 dx &\leqq \int |\phi(x, y_1)|^2 dx \\ &\quad - \int \left|\int \phi(x, y)\psi(y)\,dy\right|^2 dx.\end{aligned} \tag{2.06}$$

In all these lemmas, the limits of integration when left blank may be given any suitable values.

To establish Lemma A, we notice that

$$\begin{aligned}\int \left|\frac{1}{n}\sum_1^n \phi_k(x)\right|^2 dx &= \frac{1}{n^2}\sum_{j=1}^n\sum_{k=1}^n \int \phi_j(x)\overline{\phi_k(x)}\,dx \\ &\leqq \frac{1}{n^2}\sum_{j=1}^n\sum_{k=1}^n \left\{\int |\phi_j(x)|^2\,dx \int |\phi_k(x)|^2\,dx\right\}^{\frac{1}{2}} = A.\end{aligned} \tag{2.07}$$

As to Lemma B, we may obviously choose k_1 so that

$$R\int \overline{\phi_{k_1}(x)}\left(\frac{1}{n}\sum_{k=1}^{n}\phi_k(x)\right)dx \geqq \int\left|\frac{1}{n}\sum_{k=1}^{n}\phi_k(x)\right|^2 dx. \tag{2.08}$$

Otherwise we should have

$$R\frac{1}{n}\sum_{j=1}^{n}\int \overline{\phi_j(x)}\left(\frac{1}{n}\sum_{k=1}^{n}\phi_k(x)\right)dx < \int\left|\frac{1}{n}\sum_{k=1}^{n}\phi_k(x)\right|^2 dx, \tag{2.09}$$

and this is a contradiction. Then

$$\begin{aligned}\int\left|\phi_{k_1}(x)-\frac{1}{n}\sum_{k=1}^{n}\phi_k(x)\right|^2 dx &= \int|\phi_{k_1}(x)|^2 dx - 2R\int\overline{\phi_{k_1}(x)}\left(\frac{1}{n}\sum_{k=1}^{n}\phi_k(x)\right)dx \\ &\quad + \int\left|\frac{1}{n}\sum_{k=1}^{n}\phi_k(x)\right|^2 dx \\ &\leqq \int|\phi_{k_1}(x)|^2 dx - \int\left|\frac{1}{n}\sum_{k=1}^{n}\phi_k(x)\right|^2 dx.\end{aligned} \tag{2.10}$$

Lemmas A′ and B′ are proved in an exactly analogous manner.

We now turn to the proof of Theorem I. Let us put

$$f_{(k)}(P) = \frac{1}{2^k}\sum_{j=0}^{2^k-1} f(T^jP). \tag{2.11}$$

We then have for $k > m$

$$f_{(k)}(P) = \frac{1}{2^{k-m}}\sum_{j=0}^{2^{k-m}-1} f_{(m)}(T^{2^m j}P). \tag{2.12}$$

Thus by Lemma A, the numbers $\int_S |f_{(k)}(P)|^2\, dV_P$ form a decreasing sequence of positive terms, so that

$$F = \lim_{k\to\infty}\int_S |f_{(k)}(P)|^2\, dV_P \tag{2.13}$$

exists and is non-negative. By Lemma B, if $k > m > 0$, there exists an integer j in the range $0, 2^{k-m}-1$ such that

$$\begin{aligned}\int_S |f_{(m)}(T^{2^m j}P) - f_{(k)}(P)|^2\, dV_P &\leqq \int_S |f_{(m)}(T^{2^m j}P)|^2\, dV_P - \int_S |f_{(k)}(P)|^2\, dV_P \\ &\leqq \int_S |f_{(m)}(P)|^2\, dV_P - F,\end{aligned} \tag{2.14}$$

since T preserves measure. Also

$$\int_S |f_{(m)}(T^r P) - f_{(k)}(T^{r-2^m i} P)|^2 dV_P \leqq \int_S |f_{(m)}(P)|^2 dV_P - F. \tag{2.15}$$

Now let m_μ be so large that

$$\int_S |f_{(m_\mu)}(P)|^2 dV_P < 2^{-\mu} + F. \tag{2.16}$$

It will follow that there exists a sequence of numbers ν_μ such that

$$\int_S |f_{(m_\mu)}(T^{\nu_\mu} P) - f_{(m_{\mu+1})}(T^{\nu_{\mu+1}} P)|^2 dV_P < 2^{1-\mu}. \tag{2.17}$$

Then if $\mu > \lambda$,

$$\int_S |f_{(m_\mu)}(T^{\nu_\mu} P) - f_{(m_\lambda)}(T^{\nu_\lambda} P)|^2 dV_P \leqq \frac{2^{\frac{1}{2}(1-\lambda)}}{2^{\frac{1}{2}} - 1}. \tag{2.18}$$

Thus, by the Riesz-Fischer theorem, there exists a function $f_1(P)$ such that

$$\lim_{\mu \to \infty} \int_S |f_1(P) - f_{(m_\mu)}(T^{\nu_\mu} P)|^2 dV_P = 0. \tag{2.19}$$

Again,

$$\lim_{\mu \to \infty} \int_S |f_1(TP) - f_{(m_\mu)}(T^{\nu_\mu + 1} P)|^2 dV_P = 0. \tag{2.20}$$

Now,

$$\begin{aligned} &\int_S |f_{(m_\mu)}(T^{\nu_\mu} P) - f_{(m_\mu)}(T^{\nu_\mu + 1} P)|^2 dV_P \\ &\qquad = \int_S |2^{-m_\mu} f(T^{\nu_\mu} P) - 2^{-m_\mu} f(T^{\nu_\mu + 2^{m_\mu}} P)|^2 dV_P \\ &\qquad \leqq 2^{2-2m_\mu} \int_S |f(P)|^2 dV_P. \end{aligned} \tag{2.21}$$

Thus

$$\lim_{\mu \to \infty} \int_S |f_{(m_\mu)}(T^{\nu_\mu} P) - f_{(m_\mu)}(T^{\nu_\mu + 1} P)|^2 dV_P = 0, \tag{2.22}$$

and

$$\int_S |f_1(P) - f_1(TP)|^2 dV_P = 0. \tag{2.23}$$

Hence for all n

$$f_1(P) = f_1(TP) = f_1(T^n P). \tag{2.24}$$

Now let us put

$$f_2(P) = f(P) - f_1(P). \tag{2.25}$$

Then there exists a sequence $\{m_\mu\}$ such that

$$\lim_{\mu\to\infty}\int_S |f_{2(m_\mu)}(P)|^2\,dV_P = 0. \tag{2.26}$$

However, as in (2.13),

$$\lim_{k\to\infty}\int_S |f_{2(k)}(P)|^2\,dV_P \tag{2.27}$$

exists. Thus

$$\lim_{k\to\infty}\int_S |f_{(k)}(P) - f_1(P)|^2\,dV_P = 0. \tag{2.28}$$

Now, let the integer N have the development

$$N = 2^\nu + \sum_{\mu=0}^{\nu-1} a_\mu 2^\mu \qquad (a_\mu = 0 \text{ or } 1;\ a_\nu = 1) \tag{2.29}$$

in the binary scale. Then

$$\int_S \left| f_1(P) - \frac{1}{N}\sum_{k=0}^{N-1} f(T^k P)\right|^2 dV_P \leqq \left\{\sum_{\mu=0}^{\nu}\frac{a_\mu 2^\mu}{N}\left\{\int |f_1(P) - f_{(\mu)}(P)|^2\,dV_P\right\}^{\frac{1}{2}}\right\}^2. \tag{2.30}$$

This establishes (1.01) and Theorem I.

Theorem I′ may be established in a like manner. We shall proceed to the more general case of Theorem I″. Here we shall prove in the first instance that there exists a function $f_1(P)$ of L^2 such that

$$\lim_{\Lambda\to\infty}\int_S \left| f_1(P) - \frac{1}{(2\pi\Lambda)^{\frac{1}{2}n}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(T_1^{\lambda_1}\cdots T_n^{\lambda_n}P)\cdot\exp\left(-\sum_1^n\frac{\lambda_k^2}{2\Lambda}\right)d\lambda_1\cdots d\lambda_n\right|^2 dV_P = 0. \tag{2.31}$$

Let us notice that

$$\frac{1}{[2\pi(\Lambda_1+\Lambda_2)]^{\frac{1}{2}n}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(T_1^{\lambda_1}\cdots T_n^{\lambda_n}P)\exp\left(-\sum_1^n\frac{\lambda_k^2}{2(\Lambda_1+\Lambda_2)}\right)d\lambda_1 \cdots d\lambda_n = \frac{1}{(2\pi\Lambda_1)^{\frac{1}{2}n}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left(-\sum_1^n\frac{\lambda_k^2}{2\Lambda_1}\right)d\lambda_1\cdots d\lambda_n \frac{1}{(2\pi\Lambda_2)^{\frac{1}{2}n}} \cdot\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(T_1^{\lambda_1+\mu_1}\cdots T_n^{\lambda_n+\mu_n}P)\exp\left(-\sum_1^n\frac{\mu_k^2}{2\Lambda_2}\right)d\mu_1\cdots d\mu_n. \tag{2.32}$$

By Lemma A′,

$$\int_S \left|\frac{1}{(2\pi\Lambda)^{\frac{1}{2}n}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(T_1^{\lambda_1}\cdots T_n^{\lambda_n}P)\exp\left(-\sum_1^n\frac{\lambda_k^2}{2\Lambda}\right)d\lambda_1\cdots d\lambda_n\right|^2 dV_P \tag{2.33}$$

is monotone decreasing in Λ and has a non-negative limit F. Thus by Lemma B′, if $\Lambda > \Lambda_1 > 0$, there exists a set $(\mu_1, \cdots, \mu_n)$ such that

$$\begin{aligned}\int_S \Bigg| &\frac{1}{(2\pi\Lambda)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P) \exp\left(-\sum_1^n \frac{\lambda_k^2}{2\Lambda}\right) d\lambda_1 \cdots d\lambda_n \\ &- \frac{1}{(2\pi\Lambda_1)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(T_1^{\lambda_1+\mu_1} \cdots T_n^{\lambda_n+\mu_n} P) \exp\left(-\sum_1^n \frac{\lambda_k^2}{2\Lambda_1}\right) d\lambda_1 \\ &\cdots d\lambda_n \Bigg|^2 dV_P \leqq \int_S \Bigg| \frac{1}{(2\pi\Lambda_1)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P) \\ &\cdot \exp\left(-\sum_1^n \frac{\lambda_k^2}{2\Lambda_1}\right) d\lambda_1 \cdots d\lambda_n \Bigg|^2 dV_P - B.\end{aligned} \tag{2.34}$$

As in the proof of (2.19), we may choose the vectors $(\mu_{\nu_1}, \cdots, \mu_{\nu_n})$ and the numbers Λ_ν in such a way that the sequence

$$\frac{1}{(2\pi\Lambda_\nu)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(T_1^{\lambda_1+\mu_{\nu_1}} \cdots T_n^{\lambda_n+\mu_{\nu_n}} P) \exp\left(-\sum_1^n \frac{\lambda_k^2}{2\Lambda_\nu}\right) d\lambda_1 \cdots d\lambda_n \tag{2.35}$$

converges in the mean in the L^2 sense to a function $f_1(P)$.

The argument corresponding to (2.21)–(2.23) proceeds as follows. We have

$$\begin{aligned}\lim_{\Lambda_\nu \to \infty} \frac{1}{(2\pi\Lambda_\nu)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Bigg| &\exp\left(-\sum_1^n \frac{(\lambda_k + \mu_{\nu_k} + \mu_k)^2}{2\Lambda_\nu}\right) \\ &- \exp\left(-\sum_1^n \frac{(\lambda_k + \mu_{\nu_k})^2}{2\Lambda_\nu}\right) \Bigg| d\lambda_1 \cdots d\lambda_n = 0.\end{aligned} \tag{2.36}$$

Again,

$$\begin{aligned}&\frac{1}{(2\pi\Lambda_\nu)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \{f(T_1^{\lambda_1+\mu_{\nu_1}+\mu_1} \cdots T_n^{\lambda_n+\mu_{\nu_n}+\mu_n} P) \\ &\qquad - f(T_1^{\lambda_1+\mu_{\nu_1}} \cdots T_n^{\lambda_n+\mu_{\nu_n}} P)\} \exp\left(-\sum_1^n \frac{\lambda_k^2}{2\Lambda_\nu}\right) d\lambda_1 \cdots d\lambda_n \\ &= \frac{1}{(2\pi\Lambda_\nu)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int_{\infty}^{\infty} f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P) \left\{\exp\left(-\sum_1^n \frac{(\lambda_k + \mu_{\nu_k} + \mu_k)^2}{2\Lambda_\nu}\right)\right. \\ &\qquad \left. - \exp\left(-\sum_1^n \frac{(\lambda_k + \mu_{\nu_k})^2}{2\Lambda_\nu}\right)\right\} d\lambda_1 \cdots d\lambda_n.\end{aligned} \tag{2.37}$$

Thus by Lemma A′,

$$\begin{aligned}\lim_{\Lambda_\nu \to \infty} \int_S \Bigg| &\frac{1}{(2\pi\Lambda_\nu)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(T_1^{\lambda_1+\mu_{\nu_1}+\mu_1} \cdots T_n^{\lambda_n+\mu_{\nu_n}+\mu_n} P) \\ &\cdot \exp\left(-\sum_1^n \frac{\lambda_k^2}{2\Lambda_\nu}\right) d\lambda_1 \cdots d\lambda_n \\ &- \frac{1}{(2\pi\Lambda_\nu)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(T_1^{\lambda_1+\mu_{\nu_1}} \cdots T_n^{\lambda_n+\mu_{\nu_n}} P) \\ &\cdot \exp\left(-\sum_1^n \frac{\lambda_k^2}{2\Lambda_\nu}\right) d\lambda_1 \cdots d\lambda_n \Bigg|^2 dV_P = 0.\end{aligned} \tag{2.38}$$

The argument of (2.23)–(2.28) is duplicated without substantial change.

To proceed from (2.31) to (1.14), let us notice that if $S_\Lambda(\lambda_1, \cdots, \lambda_n)$ is the characteristic function of the sphere $\lambda_1^2 + \lambda_2^2 + \cdots + \lambda_n^2 \leqq \Lambda^2$, we have

$$\lim_{\Lambda\to\infty} \frac{1}{V(\Lambda)} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Big| S_\Lambda(\lambda_1, \cdots, \lambda_n) - \frac{1}{(2\pi\Lambda_1)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} S_\Lambda(\lambda_1 + \mu_1, \cdots, \lambda_n + \mu_n) \cdot \exp\left(-\sum_1^n \frac{\mu_k^2}{2\Lambda_1}\right) d\mu_1 \cdots d\mu_n \Big|^2 d\lambda_1 \cdots d\lambda_n = 0. \tag{2.39}$$

Thus, by Lemma A′,

$$\lim_{\Lambda\to\infty} \int_S dV_P \Big| \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \{f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P) - f_1(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P)\} \frac{1}{V(\Lambda)} \Big\{ S_\Lambda(\lambda_1, \cdots, \lambda_n) - \frac{1}{(2\pi\Lambda_1)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} S_\Lambda(\lambda_1 + \mu_1, \cdots, \lambda_n + \mu_n) \cdot \exp\left(-\sum_1^n \frac{\mu_k^2}{2\Lambda_1}\right) d\mu_1 \cdots d\mu_n \Big\} d\lambda_1 \cdots d\lambda_n \Big|^2 = 0. \tag{2.40}$$

Hence by (2.31)

$$\overline{\lim_{\Lambda\to\infty}} \int_S \Big| f_1(P) - \frac{1}{V(\Lambda)} \int \cdots \int_{\lambda_1^2 + \cdots + \lambda_n^2 \leqq \Lambda^2} f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P)\, d\lambda_1 \cdots d\lambda_n \Big|^2 dV_P \leqq \overline{\lim_{\Lambda_1\to\infty}} \int_S \Big| f_1(P) - \frac{1}{(2\pi\Lambda_1)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P) \cdot \exp\left(-\sum_1^n \frac{\lambda_k^2}{2\Lambda_1}\right) d\lambda_1 \cdots d\lambda_n \Big|^2 dV_P = 0. \tag{2.41}$$

Theorems I, I′, and I″ deal with functions of the class L^2. From these, similar theorems arise concerning functions of the class L. If in the hypothesis of any of these theorems we assume $f(P)$ only to belong to L, we shall show that (1.01) may be replaced by

$$\lim_{N\to\infty} \int_S \Big| f_1(P) - \frac{1}{N+1} \sum_{n=0}^{N} f(T^n P) \Big| dV_P = 0; \tag{2.42}$$

(1.04) by

$$\lim_{N\to\infty} \int_S \Big| f_1(P) - \frac{1}{N} \int_0^N f(T^\lambda P)\, d\lambda \Big| dV_P = 0; \tag{2.43}$$

and (1.14) by

$$\lim_{\Lambda\to\infty}\int_S \left| f_1(P) - \frac{1}{V(\Lambda)} \int\cdots\int_{\lambda_1^2+\cdots+\lambda_n^2\leq\Lambda^2} f(T_1^{\lambda_1}\cdots T_n^{\lambda_n}P)\,d\lambda_1\cdots d\lambda_n \right| dV_P = 0. \tag{2.44}$$

As all the proofs are exactly alike, we shall establish (2.42). We may approximate in the L sense to $f(P)$ with any desired degree of accuracy by a bounded function $g(P)$, which obviously belongs to L^2. Next we form $g_1(P)$ as we have formed $f_1(P)$ in the L^2 case. By the Schwarz inequality,

$$\lim_{N\to\infty}\int_S \left| g_1(P) - \frac{1}{N+1}\sum_{n=0}^{N} g(T^nP) \right| dV_P = 0. \tag{2.45}$$

Again

$$\int_S \left| \frac{1}{N+1}\sum_0^N f(T^nP) - \frac{1}{N+1}\sum_0^N g(T^nP) \right| dV_P \leqq \int_S |f(P)-g(P)|\,dV_P. \tag{2.46}$$

Thus

$$\overline{\lim_{M,N\to\infty}}\int_S \left| \frac{1}{N+1}\sum_0^N f(T^nP) - \frac{1}{M+1}\sum_0^N f(T^nP) \right| dV_P \leqq 2\int_S |f(P)-g(P)|\,dV_P. \tag{2.47}$$

Since $g(P)$ converges in the mean in the L sense to $f(P)$, the right side of (2.47) tends to 0, so that we may replace it by 0. If we now use the L form of the Riesz-Fischer theorem, (2.42) follows at once.

3. We now turn to the proofs of Theorems IV and IV′. For Theorem IV′ we need a lemma of the Vitali type:[5]

LEMMA C′. *Let S be a set of points of finite outer measure in n-space. Let each point P of S be the center of a sphere $\Sigma(P)$. Then there exists a finite set of spheres $\Sigma(P)$, non-overlapping, and of total volume exceeding $A_n m_e(S)$, where A_n is a positive absolute constant depending only on the number n of dimensions.*

To establish this lemma, let us notice that it is trivially true unless the radii of the spheres $\Sigma(P)$ have a finite upper bound. If they have such a bound, let it be B. Let S_ν be that subset of S for which the spheres $\Sigma(P)$ have radii $\leqq 2^{1-\nu}B$ and $>2^{-\nu}B$. Let $S_1' = S_1$, and let us define by mathematical induction S_ν' as that part of S_ν which is more remote than $5B$ from any point of S_1', more remote than $5B\cdot 2^{-1}$ from any point of S_2', and so on, being more remote than $5B\cdot 2^{-\nu+2}$ from any point of $S_{\nu-1}'$.

Let us lay upon S_1 a mesh of cubes with side $5B$, and if any cube of this mesh contain a point of S_1, let us pick out one of these points, P_m. Then the

[5] Cf. E. W. Hobson, *The Theory of Functions of a Real Variable*, vol. 1, Cambridge, 1927, §136. Cf. also Titchmarsh, *Theory of Functions*, §11.41.

sum of the volumes of the spheres $\Sigma(P_m)$ will be more than a fixed fraction of the sum of the meshes of the cubes corresponding to them, and hence more than a fixed fraction of the volumes of these cubes, together with all adjacent cubes. None of these spheres can overlap a sphere whose center is not in an adjacent cube, and hence each can overlap at most a fixed finite number of spheres. Thus if we enumerate these spheres in any order, and discard in turn every sphere overlapping a sphere which we have not already discarded, we shall ultimately obtain a finite set of non-overlapping spheres whose total volume will exceed a fixed submultiple A_n' of the volume of all the cubes containing points of S_1, together with all adjacent cubes. These cubes contain every point of S_1, together with all the points which are in any S_r, but fail to be included in S_r' by virtue of their vicinity to points of S_1.

We now cover S_2' with a mesh of half the linear magnitude, and proceed as above. We shall find a finite set of spheres with centers at points of S_2', obviously not overlapping any sphere of the previous set, not overlapping one another, and with a total volume exceeding A_n' multiplied by the total volume of all the cubes containing S_2', together with all adjacent cubes. These cubes together with the similar set obtained from S_1 contain S_1, S_2, and all the points in later S_r, but are excluded from the corresponding S_r' on account of their propinquity to S_1 or S_2'.

By continuing this process, we obtain a denumerable set of spheres $\Sigma(P)$, exceeding in volume $A_n' m_e(S)$ and non-overlapping. We may clearly take a finite number of these with total volume exceeding $A_n m_e(S)$, if $A_n < A_n'$.

In the one-dimensional case, Lemma C′ is replaced by the somewhat more precise

LEMMA C. *Let S be a set of points of finite measure on the line. Let every point P of S be the right-hand terminus of an interval $I(P)$. Then there exists a finite subset of intervals $I(P)$, non-overlapping, and of total length exceeding $(1 - \epsilon)m_e(S)$, where ϵ is any positive number.*

This lemma is due to Sierpinski (see footnote 5) and is completely elementary.

From Lemma C and Lemma C′, respectively, we may deduce

LEMMA D. *Let $f(x)$ belong to L on the infinite line. Then the set S_α of values of x for which*

$$\text{(3.01)} \qquad \text{l.u.b.}_{0<\Lambda<\infty} \frac{1}{\Lambda} \int_{x-\Lambda}^{x} |f(\xi)|\, d\xi > \alpha > 0$$

does not exceed $\frac{1}{\alpha} \int_{-\infty}^{\infty} |f(\xi)|\, d\xi$ in measure.

LEMMA D′. *Let $f(x_1, \cdots, x_n)$ belong to L in infinite n-space. Then the set S_α of points $(y_1, \cdots, y_n)$ for which*

$$\text{(3.02)} \qquad \text{l.u.b.}_{0<\Lambda<\infty} \frac{1}{V(\Lambda)} \int\cdots\int_{x_1^2+\cdots+x_n^2 \le \Lambda^2} |f(x_1 + y_1, \cdots, x_n + y_n)|\, dx_1 \cdots dx_n > \alpha > 0$$

does not exceed

$$\frac{1}{\alpha A_n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}|f(x_1,\cdots,x_n)|\,dx_1\cdots dx_n \tag{3.03}$$

in measure.

The proof is the same in both cases. In Lemma D, we can pick out a finite set of non-overlapping intervals, of total length exceeding $(1-\epsilon)m(S_\alpha)$ and such that the integral of $f(x)$ over each interval exceeds α multiplied by its length. Thus

$$\alpha(1-\epsilon)m(S_\alpha)\leqq\int_\Sigma|f(\xi)|\,d\xi, \tag{3.04}$$

where Σ is a finite set of non-overlapping intervals.

The discrete analogue of Lemma D is

LEMMA E. *Let* $\sum_{-\infty}^{\infty}|A_\nu|<\infty$. *Then the number of values of* ν *for which*

$$\text{l.u.b.}_{0<\mu<\infty}\frac{1}{\mu}\sum_{k=\nu-\mu+1}^{k=\nu}|A_k|>\alpha>0 \tag{3.05}$$

does not exceed $\frac{1}{\alpha}\sum_{-\infty}^{\infty}|A_k|$.

Here the proof goes exactly as in Lemma D. The discrete equivalent of Lemma C is trivial.

We proceed to the proof of Theorem IV. It is clear that, for a particular P, the number of values of ν on $(0, N)$ for which

$$\text{l.u.b.}_{0<\mu\leq\nu}\frac{1}{\mu}\sum_{k=\nu-\mu+1}^{k=\nu}f(T^{-k}P)>\alpha>0 \tag{3.06}$$

does not exceed

$$\frac{1}{\alpha}\sum_{0}^{N}f(T^{-k}P). \tag{3.07}$$

If we now integrate over S with respect to P, we see that the sum of the measures of the sets of values of P for which

$$\text{l.u.b.}_{0<\mu\leq\nu}\frac{1}{\mu}\sum_{k=\nu-\mu+1}^{k=\nu}f(T^{-k}P)>\alpha>0\qquad(\nu=1,2,\cdots,N) \tag{3.08}$$

will not exceed

$$\frac{N+1}{\alpha}\int_S f(P)\,dV_P. \tag{3.09}$$

Since $T^{-\nu}$ is an equimeasure transformation, we get

$$\frac{1}{N}\sum_{\nu=1}^{N}\Big(\text{measure of set of values of } P \text{ for which } \text{l.u.b.}_{0<\mu\leq\nu}\frac{1}{\mu}\sum_{k=0}^{k=\mu-1}f(T^kP)>\alpha>0\Big)\leqq\frac{1}{\alpha}\int_S f(P)\,dV_P. \tag{3.10}$$

We now use the simple lemma that if B_ν is an increasing sequence for which

$$\frac{1}{N+1}\sum_{\nu=1}^{N} B_\nu \leqq B \qquad \text{for all } N, \tag{3.11}$$

then

$$B_\nu \leqq B \qquad \text{for all } \nu. \tag{3.12}$$

Thus for all values of ν,

$$\left(\text{measure of set of values of } P \text{ for which } \underset{0<\mu\leqq\nu}{\text{l.u.b.}} \frac{1}{\mu}\sum_{k=0}^{\mu-1} f(T^k P) > \alpha > 0\right) \leqq \frac{1}{\alpha}\int_S f(P)\, dV_P, \tag{3.13}$$

which gives us

$$\left(\text{measure of set of all values of } P \text{ for which } \underset{0<\mu<\infty}{\text{l.u.b.}} \frac{1}{\mu}\sum_{k=0}^{\mu-1} f(T^k P) > \alpha > 0\right) \leqq \frac{1}{\alpha}\int_S f(P)\, dV_P. \tag{3.14}$$

Let us now turn to Theorem IV′. In Lemma D′, let

$$f(x_1, \cdots, x_n) = \begin{cases} f(T_1^{x_1} \cdots T_n^{x_n} P) & \text{if } x_1^2 + \cdots + x_n^2 \leqq \Lambda^2; \\ 0 & \text{otherwise.} \end{cases} \tag{3.15}$$

Then if we integrate with respect to P over S, we see that in the product space of $(\mu_1, \cdots, \mu_n)$ and P, the measure of the set of points for which

$$\underset{0<\Lambda<\infty}{\text{l.u.b.}} \frac{1}{V(\Lambda)} \int_{\substack{\lambda_1^2+\cdots+\lambda_n^2\leqq\Lambda^2 \\ (\lambda_1+\mu_1)^2+\cdots+(\lambda_n+\mu_n)^2\leqq\Lambda^2}} \cdots \int f(T_1^{\lambda_1+\mu_1} \cdots T_n^{\lambda_n+\mu_n} P)\, d\lambda_1 \cdots d\lambda_n > \alpha > 0 \tag{3.16}$$

does not exceed

$$\frac{1}{\alpha A_n}\int_S dV_P \int_{\lambda_1^2+\cdots+\lambda_n^2\leqq\Lambda^2} \cdots \int f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P)\, d\lambda_1 \cdots d\lambda_n = \frac{V(\Lambda)}{\alpha A_n}\int_S f(P)\, dV_P. \tag{3.17}$$

Since $T_1^{\mu_1} \cdots T_n^{\mu_n}$ is an equimeasure transformation, we may replace the measure of the set for which (3.16) holds by

$$\int \cdots \int d\mu_1 \cdots d\mu_n \left\{\text{measure of the set of values of } P \text{ for which } \underset{0<\Lambda<\infty}{\text{l.u.b.}} \frac{1}{V(\Lambda)} \int_{\substack{\lambda_1^2+\cdots+\lambda_n^2\leqq\Lambda^2 \\ (\lambda_1+\mu_1)^2+\cdots+(\lambda_n+\mu_n)^2\leqq\Lambda^2}} \cdots \int f(T_1^{\lambda_1} \cdots T_n^{\lambda_n} P)\, d\lambda_1 \cdots d\lambda_n > \alpha > 0\right\}. \tag{3.18}$$

A fortiori,

$$\int\cdots\int_{\mu_1^2+\cdots+\mu_n^2\leq \frac{1}{4}\Lambda^2} d\mu_1\cdots d\mu_n \Big\{\text{measure of set of values of } P \text{ for which}$$

$$\text{(3.19)}\qquad \text{l.u.b.}_{0<\Lambda\leq \frac{1}{4}\Lambda}\frac{1}{V(\Lambda)}\int\cdots\int_{\lambda_1^2+\cdots+\lambda_n^2\leq\Lambda^2} f(T_1^{\lambda_1}\cdots T_n^{\lambda_n}P)\,d\lambda_1\cdots d\lambda_n > \alpha > 0\Big\}$$

$$\leqq \frac{V(\Lambda)}{\alpha A_n}\int_S f(P)\,dV_P.$$

Thus

$$\Big\{\text{measure of set of values of } P \text{ for which}$$

$$\text{(3.20)}\qquad \text{l.u.b.}_{0<\Lambda\leq \frac{1}{4}\Lambda}\frac{1}{V(\Lambda)}\int\cdots\int_{\lambda_1^2+\cdots+\lambda_n^2\leq\Lambda^2} f(T_1^{\lambda_1}\cdots T_n^{\lambda_n}P)\,d\lambda_1\cdots d\lambda_n > \alpha > 0\Big\}$$

$$\leqq \frac{2^n}{\alpha A_n}\int_S f(P)\,dV_P.$$

If we now put $B_n = A_n 2^{-n}$, Theorem IV′ follows at once.

Theorems IV and IV′ have the secondary conclusions (1.11) and (1.19), respectively. To obtain these, let us introduce the function $h(P)$ which is $f(P)$ when this exceeds $\frac{1}{2}\alpha$ and is 0 otherwise. Clearly

$$\text{(3.21)}\qquad f^*(P) \leqq h^*(P) + \tfrac{1}{2}\alpha.$$

Thus in Theorem IV

$$\text{(3.22)}\qquad \begin{aligned}&\{\text{measure of set of values of } P \text{ for which } f^*(P) > \alpha\}\\ &\leqq \{\text{measure of set of values of } P \text{ for which } h^*(P) > \tfrac{1}{2}\alpha\}\\ &\leqq \frac{2}{\alpha}\int_S h(P)\,dV_P = \frac{2}{\alpha}\int_{f(P)\geq\frac{1}{2}\alpha} f(P)\,dV_P.\end{aligned}$$

This yields (1.11), and (1.19) may be obtained in the identical manner. These are similar to certain results of Hardy and Littlewood, but are somewhat weaker.

Let us now turn to the proof of Theorem II. By (2.42), we may choose N so large that

$$\text{(3.23)}\qquad \int_S \left| f_1(P) - \frac{1}{N+1}\sum_{n=0}^{N} f(T^n P)\right| dV_P < \epsilon^2.$$

Then by Theorem IV, except for a set of values of P of measure not exceeding ϵ, we shall have

$$\text{(3.24)}\qquad \text{l.u.b.}_{0<A<\infty}\left|\frac{1}{A+1}\sum_{n=0}^{A}\left[f_1(T^n P) - \frac{1}{N+1}\sum_{m=0}^{N} f(T^{m+n}P)\right]\right| < \epsilon.$$

We have, for large A,

$$\begin{aligned} \frac{1}{A+1}\sum_{N+1}^{A} f(T^n P) &\leqq \frac{1}{A+1}\sum_{n=0}^{A}\frac{1}{N+1}\sum_{m=0}^{N} f(T^{m+n}P) \\ &\leqq \frac{1}{A+1}\sum_{0}^{A+N} f(T^n P). \end{aligned} \tag{3.25}$$

Thus, except for a set of values of P of measure not exceeding ϵ, we have

$$\overline{\lim_{A\to\infty}}\left| f_1(P) - \frac{1}{A+1}\sum_{n=0}^{A} f(T^n P)\right| < \epsilon. \tag{3.26}$$

Since ϵ can be given consecutively values from a convergent series, we see that, except over a set whose measure does not exceed the remainder of this series, we have uniformly

$$f_1(P) = \lim_{A\to\infty}\frac{1}{A+1}\sum_{n=0}^{A} f(T^n P). \tag{3.27}$$

This establishes the Birkhoff ergodic theorem. The parallel Theorems II′ and II″ are proved in exactly the same way.

Theorem III and its n-dimensional analogue, which is the particular case of Theorem III″ which we obtain when the group reduces to the n-dimensional translation group, may be proved on the basis of Theorems IV and IV′, or even on the basis of the Lemmas D and D′. Every function of class L is the sum of a continuous bounded function and a function of small L-norm. Let $f(x)$ be the sum of $\phi(x)$ and $\psi(x)$, where $\phi(x)$ is continuous and $\int_{-\infty}^{\infty} |\psi(x)|\,dx < \frac{1}{2}\epsilon^2$. Then

$$\begin{aligned} \overline{\lim_{\eta\to 0}}\left| f(x) - \frac{1}{\eta}\int_x^{x+\eta} f(\xi)\,d\xi\right| &= \overline{\lim_{\eta\to 0}}\left|\psi(x) - \frac{1}{\eta}\int_x^{x+\eta}\psi(\xi)\,d\xi\right| \\ &\leqq |\psi(x)| + \psi^*(x), \end{aligned} \tag{3.28}$$

where

$$\psi^*(x) = \text{l.u.b.}_{0<\eta<\infty}\frac{1}{\eta}\int_x^{x+\eta}|\psi(\xi)|\,d\xi. \tag{3.29}$$

By Lemma D, the set of points for which $\psi^*(x) > \epsilon$ has a measure not exceeding $\frac{1}{2}\epsilon$, and the same is clearly true of $|\psi(x)|$. Thus, except for a set of points of measure not exceeding ϵ, we have

$$\overline{\lim_{\eta\to 0}}\left| f(x) - \frac{1}{\eta}\int_x^{x+\eta} f(\xi)\,d\xi\right| \leqq \epsilon. \tag{3.30}$$

The proof now proceeds as in Theorem II. Theorem III″, as far as the translation group is concerned, is proved in the same way.

To prove Theorem III′ and the general case of Theorem III″, let us note that, for a fixed P, the set of values of λ for which we fail to have

$$f(T^{\lambda}P) = \lim_{\eta \to 0} \frac{1}{\eta} \int_{\lambda}^{\lambda+\eta} f(T^{\xi}P)\, d\xi, \tag{3.31}$$

or the set of values of $(\lambda_1, \cdots, \lambda_n)$ for which we fail to have

$$f(T_1^{\lambda_1} \cdots T_n^{\lambda_n}P) = \lim_{\Delta \to 0} \frac{1}{V(\Delta)} \int \cdots \int_{\mu_1^2+\cdots+\mu_n^2 \leq \Delta^2} f(T_1^{\lambda_1+\mu_1} \cdots T_n^{\lambda_n+\mu_n}P)\, d\mu_1 \cdots d\mu_n \tag{3.32}$$

are sets of zero measure. Thus in the product space of these spaces and P, the corresponding sets of values are of zero measure. That is, for almost all values of λ or $(\lambda_1, \cdots, \lambda_n)$, the set of points P for which we fail to have (3.31) or (3.32) valid is of zero measure. Therefore, for at least one of these values, the set is of zero measure. The result follows by an equimeasure transformation.

We finally come to Theorem V. Let $M(x)$ be the measure of the set of points over which $f(P) > x$, and let $M^*(x)$ be the measure of the set of points over which $f^*(P) \geqq x$. We shall have

$$\begin{aligned} \int_S s(f(P))\, dV_P &= -\int_0^{\infty} s(x)\, dM(x); \\ \int_S s(f^*(P))\, dV_P &= -\int_0^{\infty} s(x)\, dM^*(x). \end{aligned} \tag{3.33}$$

By Theorem IV, we have formally for $\nu > 0$

$$\begin{aligned} \int_0^{\infty} m^*(x)x^{\nu}\, dx &\leqq -\text{const.} \int_0^{\infty} x^{\nu-1}\, dx \int_x^{\infty} y dM(y) \\ &= -\text{const.} \int_0^{\infty} y dm(y) \int_0^{y} x^{\nu-1}\, dx \\ &= -\text{const.} \int_0^{\infty} y^{\nu+1}\, dM(y). \end{aligned} \tag{3.34}$$

Thus if $\int_0^{\infty} y^{\nu+1}\, dM(y)$ is bounded, it will follow that

$$0 = \lim_{\xi \to \infty} \int_{\xi}^{2\xi} m^*(x)x^{\nu}\, dx = \lim_{\xi \to \infty} M^*(2\xi)\xi^{\nu+1}. \tag{3.35}$$

We may hence integrate by parts in (3.34) and obtain

$$-\int_0^{\infty} x^{\nu+1}\, dm^*(x) \leqq -\text{const.} \int_0^{\infty} y^{\nu+1}\, dM(y). \tag{3.36}$$

Similarly,

$$\int_1^\infty m^*(x)\,dx \leqq -\text{const.} \int_1^\infty y\,dM(y) \int_1^y \frac{dx}{x}$$

$$= -\text{const.} \int_1^\infty y \log y\,dM(y). \tag{3.37}$$

As before, if $\int f(P) \log^+ f(P)\, dV_P$ is bounded,

$$-\int_1^\infty x\,dm^*(x) \leqq -\text{const.} \int_1^\infty y \log y\,dM(y). \tag{3.38}$$

These theorems, apart from the less accurate constants, are the same as theorems given by Hardy, Littlewood, and Pólya (loc. cit.).

Massachusetts Institute of Technology.

THE MEAN ERGODIC THEOREM

By Garrett Birkhoff

1. **Introduction.** Von Neumann's Mean Ergodic Theorem will be given an elementary proof below in the following generalized form:

Theorem. *Let X be any uniformly convex Banach space, and T any linear operator on X such that $|xT| \leqq |x|$ for all x. Then the "n-th means"*

$$g_n = \frac{1}{n}(f + fT + \cdots + fT^{n-1})$$

of the transforms fT^k of any $f \in X$, under powers of T, converge to a limit.

Uniform convexity is understood in the sense of J. A. Clarkson to mean that $|x| \leqq |y| \leqq 1$ and $|x - y| \geqq \epsilon$ imply $|\frac{1}{2}(x + y)| \leqq |y| - u(\epsilon)$, where $u(\epsilon) > 0$ if $\epsilon > 0$. Equivalently, $|x| \leqq |y| \leqq 1$ and $|y| - |\frac{1}{2}(x + y)| \leqq \epsilon$ imply $|x - y| \leqq w(\epsilon)$, where $w(\epsilon) \downarrow 0$ as $\epsilon \downarrow 0$. Clarkson gives[1] an elementary proof that the spaces L_p $(p > 1)$ are uniformly convex.

The author has had access to the following unpublished papers: F. Riesz, *Some mean ergodic theorems*; N. Wiener, *The ergodic theorem* (published in this issue of this Journal; S. Kakutani, *Weak convergence in uniformly convex spaces*; K. Yosida, *Mean ergodic theorem in Banach spaces.*

The author's proof is most closely related to Wiener's, which alone does not depend on the non-elementary theorem that the unit sphere in Hilbert space is weakly compact. It is more general and somewhat shorter than Wiener's, which applies only to isometric transformations of L_2, or Riesz', which is only sketched for L_p $(p \neq 2)$. It is much shorter than the proofs of Kakutani and Yosida, although less general.

2. **Verbal argument.** The key point is that unless g_nT^{kn} is very near g_n, $|\frac{1}{2}(g_n + g_nT^{kn})|$ must by uniform convexity be appreciably less than $|g_n|$. Hence the norm of the mean of the $\frac{1}{2}(g_n + g_nT^{kn})T^{hn}$, for $h = 0, \cdots, k - 1$, will also be appreciably less than $|g_n|$. But this mean is

$$\left|\frac{1}{2k}(g_n + g_nT + \cdots + g_nT^{2kn-n})\right| = |g_{2kn}|.$$

It follows that if $|g_n|$ is chosen sufficiently near its infimum M, then every g_nT^{kn} will be arbitrarily near g_n. Hence so will every g_{mn}, being the mean of

Received December 29, 1938.

[1] J. A. Clarkson, *Uniformly convex spaces*, Trans. Am. Math. Soc., vol. 40(1936), pp. 396–414. In Hilbert space, $|x + y|^2 + |x - y|^2 = 2(|x|^2 + |y|^2)$, whence if $|x| \leqq |y|$, $|\frac{1}{2}(x + y)|^2 \leqq |y|^2 - |\frac{1}{2}(x - y)|^2$, and $|\frac{1}{2}(x + y)| \leqq |y| - \frac{1}{4}|x - y|^2$ if $|y| \leqq 1$.

the g_nT^{kn} for $k = 0, \cdots, m-1$. But if i is large, then g_i will be very near g_{mn} for that largest integral multiple mn of n not exceeding i. Hence *every* g_i (i very large) will be near g_n. Consequently, the g_n form a Cauchy sequence.

3. **Formal proof.** First, in any linear metric space

$$\left|\frac{1}{n}(x_1 + \cdots + x_n)\right| \leqq \max |x_i|. \tag{1}$$

From this we see that $|g_nT^i| \leqq 1$ for all n and i.

Again, without loss of generality, we can assume $|f| = 1$. Now let M be the infimum of the $|g_n|$. Given $\epsilon > 0$, for some n, $|g_n| < M + \epsilon$. Let α_k denote $|g_nT^{kn} - g_n|$; then

$$|\tfrac{1}{2}(g_n + g_nT^{kn})| < M + \epsilon - u(\alpha_k) \tag{2}$$

by uniform convexity, whence

$$|\tfrac{1}{2}(g_n + g_nT^{kn})T^{hn}| < M + \epsilon - u(\alpha_k) \qquad \text{for all } h,$$

and

$$|g_{2kn}| = \left|\frac{1}{k}\sum_{h=0}^{k-1} \tfrac{1}{2}(g_n + g_nT^{kn})T^{hn}\right|; \tag{3}$$

by grouping terms

$$|g_{2kn}| < M + \epsilon - u(\alpha_k)$$

from (1). But $|g_{2kn}| \geqq M$; hence $u(\alpha_k) < \epsilon$, and

$$|g_nT^{kn} - g_n| \leqq w(\epsilon) \qquad \text{for all } k. \tag{4}$$

Further, since $g_{mn} = m^{-1}(g_n + g_nT^n + \cdots + g_nT^{mn-n})$,

$$|g_{mn} - g_n| = \left|\frac{1}{m}\sum_{k=1}^{m-1}(g_nT^{kn} - g_n)\right| \leqq w(\epsilon) \tag{5}$$

from (4) and (1). Finally, since $g_{r+i} = r(r+i)^{-1}g_r + (r+i)^{-1}(fT^r + \cdots + fT^{r+i-1})$,

$$|g_{r+i} - g_r| \leqq \frac{1}{r+i}|g_r| + \frac{1}{r+i}|f| \leqq \frac{2i}{r+i}|f|. \tag{6}$$

Combining (6) with (5), we see that if $0 \leqq i < n$, then

$$|g_{mn+i} - g_n| \leqq |g_{mn+i} - g_{mn}| + |g_{mn} - g_n| < \frac{2}{m}|f| + w(\epsilon). \tag{7}$$

Hence from a certain point on, every g_j is within $\epsilon + w(\epsilon)$ of g_n, and so within $2\epsilon + 2w(\epsilon)$ of every other such g_j. But by choosing ϵ appropriately, we can make $2\epsilon + 2w(\epsilon)$ as small as we please. Consequently, the *n-th means* of f form a *Cauchy sequence*. The theorem follows directly from this.

HARVARD UNIVERSITY.

THE ANALYTIC PROLONGATION OF A MINIMAL SURFACE OVER A RECTILINEAR SEGMENT OF ITS BOUNDARY

By Jesse Douglas

1. Schwarz proved the theorem that if a minimal surface contains a straight line in its interior, then this straight line must be an axis of symmetry of the surface.[1] All proofs so far given of this theorem make essential use of the supposition of the *interior* position of the line with respect to the minimal surface.

For certain interesting applications (see §6) it is important to have stronger information. Suppose the straight line segment l is part of the *boundary* of a portion M of a minimal surface. Let M be rotated about l through 180°, producing M'. Is then M' the *analytic prolongation* of M across l?

In other words, can symmetry with respect to a straight line be employed as a principle for the analytic prolongation of a minimal surface, and not merely figure as a property of the surface when already prolonged so as to contain the straight line in its interior?

The present paper gives a proof of the validity of the stated prolongation principle. In §6 an application is presented to the construction of a minimal surface bounded by two given interlacing circles. §7 poses the general problem of the analytic prolongation of a minimal surface across an analytic arc.

2. Let H denote the *interior* of the upper half of the (u, v)-plane $(v > 0)$, while ab denotes an *open* segment $(a < u < b)$ of the axis $v = 0$. Suppose that $x(u, v)$, $y(u, v)$, $z(u, v)$ are any three functions which have the following properties:

(i) they are defined and continuous on $H + ab$;

(ii) they are harmonic in H:

$$x_{uu} + x_{vv} = 0, \qquad y_{uu} + y_{vv} = 0, \qquad z_{uu} + z_{vv} = 0; \tag{1}$$

(iii) they obey throughout H the relations

$$x_u^2 + y_u^2 + z_u^2 = x_v^2 + y_v^2 + z_v^2, \tag{2}$$

$$x_u x_v + y_u y_v + z_u z_v = 0; \tag{3}$$

(iv) for (u, v) on ab, i.e., $a < u < b$, $v = 0$, the point of coördinates

Received October 28, 1938.

[1] H. A. Schwarz, *Gesammelte Mathematische Abhandlungen*, vol. 1, p. 181. For another proof, see the tract by T. Radó, *On the Problem of Plateau*, 1933, p. 30.

$$x = x(u, v), \qquad y = y(u, v), \qquad z = z(u, v) \tag{4}$$

always lies on a given straight line l.[2]

Then the equations (4) define a minimal surface M contiguous to l. By (2), (3), M is in conformal representation on H, and our method of analytic prolongation will preserve this conformal representation of M on the (u, v)-plane, as well as the character of M as a minimal surface.

Without loss of generality we may suppose l to be the z-axis, since the conditions (1), (2), (3) are invariant under any rigid motion of the (x, y, z)-space.

Then on ab $(a < u < b)$ we have

$$x(u, 0) = 0, \qquad y(u, 0) = 0. \tag{5}$$

It follows by the well-known symmetry principle of Schwarz applied to the harmonic functions $x(u, v)$, $y(u, v)$ that the formulas

$$x(u, -v) = -x(u, v), \qquad y(u, -v) = -y(u, v) \tag{6}$$

define an analytic prolongation of these functions across the segment ab into the lower half H' $(v < 0)$ of the (u, v)-plane. Thus, $x(u, v)$ and $y(u, v)$ are regular also on ab, so that their partial derivatives x_u, x_v, y_u, y_v exist and are continuous jointly in (u, v) for

$$a < u < b, \qquad v = 0. \tag{7}$$

On account of (5), we have on the open segment ab, or (7),

$$x_u = 0, \qquad y_u = 0.$$

It follows from these remarks and from the fact that (2), (3) hold in H, that as we approach any fixed point P of ab from H—i.e.,

$$u \to u_0, \qquad v\ (>0) \to 0, \tag{8}$$

where $a < u_0 < b$—we have

$$z_u^2(u, v) - z_v^2(u, v) \to x_v^2(u_0, 0) + y_v^2(u_0, 0), \tag{9}$$

$$z_u(u, v) \cdot z_v(u, v) \to 0. \tag{10}$$

Forming then (9) squared plus four times (10) squared, we get

$$[z_u^2(u, v) + z_v^2(u, v)]^2 \to [x_v^2(u_0, 0) + y_v^2(u_0, 0)]^2,$$

and since the quantity in each bracket is non-negative, it follows that[3]

$$z_u^2(u, v) + z_v^2(u, v) \to x_v^2(u_0, 0) + y_v^2(u_0, 0). \tag{11}$$

[2] The correspondence so established between ab and a segment of l need not be one-one.

[3] If $x^2 \to a^2$, it does *not* follow necessarily that either $x \to a$ or $x \to -a$, since x may oscillate between a and $-a$ as limits of indetermination. But if, besides, x is always of the same sign as a, then $x \to a$. A good deal of the reasoning of the present paper is concerned with inferring $x \to a$ from $x^2 \to a^2$, under the actual auxiliary circumstances.

By taking the semi-sum and semi-difference of (9) and (11), we then find

$$z_u^2(u, v) \to x_v^2(u_0, 0) + y_v^2(u_0, 0), \tag{12}$$

$$z_v^2(u, v) \to 0. \tag{13}$$

The latter, of course, implies simply

$$z_v(u, v) \to 0. \tag{14}$$

3. We now prove—what is the essential point of our entire argument—*that* z_u, z_v *actually exist ON the segment* ab *and are continuous on the set* $H + ab$. Indeed, it will be shown that on ab these partial derivatives have the values which are indicated by the limiting relations (12) and (14), namely:

$$z_v(u_0, 0) = 0, \tag{15}$$

$$z_u(u_0, 0) = \pm\sqrt{x_v^2(u_0, 0) + y_v^2(u_0, 0)}. \tag{16}$$

The choice of sign in (16) for each particular value of u_0 is such as to be consistent with the stated continuity property of z_u.

Proof of (15). By the law of the mean,

$$\frac{z(u_0, h) - z(u_0, 0)}{h} = z_v(u_0, \theta h),$$

where $0 < \theta < 1$. This applies with full validity, since $z(u_0, v)$ is continuous as a function of v for $0 \leqq v \leqq h$, while $z_v(u_0, v)$ exists for $0 < v < h$. It follows by (14) (with reference to (8)) that $z_v(u_0, 0)$ exists and has the value zero, as expressed by (15).

Continuity of z_v *on* $H + ab$. This is an immediate consequence of (14) (with (8)) and (15).

Proof of (16). We have by the continuity of $z(u, v)$ that[4]

$$\frac{z(u_0 + h, \epsilon) - z(u_0, \epsilon)}{h} \to \frac{z(u_0 + h, 0) - z(u_0, 0)}{h},$$

when $\epsilon > 0$ tends to zero. By the law of the mean, then,

$$z_u(u_0 + \theta h, \epsilon) \to \frac{z(u_0 + h, 0) - z(u_0, 0)}{h}, \tag{17}$$

where θ is a function of ϵ such that $0 < \theta < 1$. We can make $\epsilon \to 0$ on a sequence ϵ_n so that θ tends to some particular value θ_0, where $0 \leqq \theta_0 \leqq 1$. Then the point $(u_0 + \theta h, \epsilon)$ tends to $(u_0 + \theta_0 h, 0)$; consequently by (12),[5]

$$z_u^2(u_0 + \theta h, \epsilon) \to x_v^2(u_0 + \theta_0 h, 0) + y_v^2(u_0 + \theta_0 h, 0).$$

[4] Throughout this reasoning, h is supposed to be so small that $u_0 + h$ still belongs to the segment ab.

[5] $u_0 + \theta_0 h$ is a point of ab, according to the preceding footnote.

Therefore, on account of the continuity of x_v, y_v on ab, it follows that as $h \to 0$

$$\left[\frac{z(u_0 + h, 0) - z(u_0, 0)}{h}\right]^2 \to x_v^2(u_0, 0) + y_v^2(u_0, 0). \tag{18}$$

If $x_v(u_0, 0) = y_v(u_0, 0) = 0$, the result (16) is immediate. We therefore lay this case aside and suppose

$$x_v^2(u_0, 0) + y_v^2(u_0, 0) > 0. \tag{19}$$

Let now $h \to 0$ on an arbitrary particular sequence h_n. It follows from (18) that the difference quotient

$$\Delta(h_n) = \frac{z(u_0 + h_n, 0) - z(u_0, 0)}{h_n}$$

either does not approach a limit,[6] or it approaches $+\sqrt{x_v^2(u_0, 0) + y_v^2(u_0, 0)}$, or it approaches $-\sqrt{x_v^2(u_0, 0) + y_v^2(u_0, 0)}$. According to (19), the values $+\sqrt{\ }$ and $-\sqrt{\ }$ are *distinct*.

Now, the total class of limits approached by the difference quotient of a continuous function, relative to a given fixed point, is never empty,[7] and never consists merely of two distinct values such as $+\sqrt{\ }$ and $-\sqrt{\ }$, but is always either a proper[8] closed interval (finite or infinite) of values, or else a unique value—in the latter case, the function is *differentiable* at the given point with a derivative equal to the unique value of $\lim \Delta(h_n)$.

It follows that $z_u(u_0, 0)$ exists and has at each point u_0 one of the values expressed by (16).

Continuity of z_u on $H + ab$. We have to prove that if (u, v) approach $(u_0, 0)$ *from H*, i.e., in accordance with (8), then

$$z_u(u, v) \to z_u(u_0, 0). \tag{20}$$

If this can be done, it will follow easily by general continuity principles that z_u is also continuous at $(u_0, 0)$ with respect to approach along ab; i.e.,

$$z_u(u, 0) \to z_u(u_0, 0) \qquad \text{when } u \to u_0. \tag{21}$$

In order to see this particular point, let $f(u, v)$ be any function[9] defined for $v > 0$, and suppose that when

$$u \to u_0, \qquad v\ (>0) \to 0, \tag{22}$$

we have for every point u_0 in the open interval ab

$$f(u, v) \to \varphi(u_0). \tag{23}$$

[6] See here footnote 3.

[7] If we include $\pm\infty$ as permissible limiting values.

[8] Distinct end-points.

[9] Not even required, for the present specific purpose, to be continuous.

Then (I say) $\varphi(u_0)$ must be a continuous function of u_0 on the open interval ab.

For let u_0' denote a particular point of ab, while

$$\rho = \text{distance of } (u, v) \text{ from } (u_0', 0).$$

Then for arbitrary $\epsilon > 0$, there exists $\delta > 0$ such that if

$$\rho < \delta, \qquad v > 0, \tag{24}$$

then

$$\varphi(u_0') - \epsilon \leqq f(u, v) \leqq \varphi(u_0') + \epsilon. \tag{25}$$

This is exactly what is expressed by (23) (with (24)).

Suppose that u_0 is any point of the interval

$$u_0' - \delta < u_0 < u_0' + \delta. \tag{26}$$

Then by (23) and (25), $\varphi(u_0)$ is the limit of a variable all of whose values belong to the interval $(\varphi(u_0') - \epsilon, \varphi(u_0') + \epsilon)$; hence, $\varphi(u_0)$ also belongs to this interval:

$$\varphi(u_0') - \epsilon \leqq \varphi(u_0) \leqq \varphi(u_0') + \epsilon. \tag{27}$$

But that (26) implies (27) is precisely the definition of the continuity of $\varphi(u_0)$ at the arbitrary particular point u_0'.

Thus, it will suffice to prove (20) for the mode of approach (22) of (u, v) to $(u_0, 0)$.

To this end, we remark first that certainly, by (12) and (16),

$$z_u^2(u, v) \to z_u^2(u_0, 0). \tag{28}$$

If $z_u(u_0, 0) = 0$, the desired result follows immediately. We therefore lay this case aside and suppose

$$z_u(u_0, 0) \neq 0. \tag{29}$$

If (u_n, v_n) denotes any sequence tending to $(u_0, 0)$, it follows from (28) that either (a) $z_u(u_n, v_n)$ tends to no limit, or (b) $z_u(u_n, v_n) \to z_u(u_0, 0)$, or (c) $z_u(u_n, v_n) \to -z_u(u_0, 0)$.

It is readily seen that if for a particular sequence (u_n', v_n') we have case (b):

$$z_u(u_n', v_n') \to z_u(u_0, 0), \tag{30}$$

then we have this case for every sequence. Otherwise, let (u_n'', v_n'') be a sequence for which we have either case (a) or (c). If (u_n'', v_n'') is in case (a), then certainly it is not in case (b); therefore, a subsequence (u_n''', v_n''') exists such that $z_u(u_n''', v_n''')$ tends to a limit different from $z_u(u_0, 0)$, and this limit can only be $-z_u(u_0, 0)$:

$$z_u(u_n''', v_n''') \to -z_u(u_0, 0). \tag{31}$$

If, on the other hand, (u_n'', v_n'') is in case (c), then we have (31) immediately for the original sequence (u_n'', v_n'').

From (30), (31), and (29), it follows that, if we neglect only a finite number of values of n at the start, $z_u(u_n', v_n')$ and $z_u(u_n''', v_n''')$ will always have opposite signs. Therefore, on each line segment joining (u_n', v_n') and (u_n''', v_n''') there will be a point (u_n^{iv}, v_n^{iv}) such that $z_u(u_n^{iv}, v_n^{iv}) = 0$—for $z_u(u, v)$ is continuous on H. Thus $z_u(u_n^{iv}, v_n^{iv}) \to 0$. But then, as we see by (29), the sequence (u_n^{iv}, v_n^{iv}) is in none of the cases (a), (b), (c), and this is impossible. Thus the proof of the proposition at the beginning of the preceding paragraph is complete.

Accordingly, we have only to show the existence of one particular sequence obeying (30). This is done as follows. By definition of derivative, we can find an arbitrary small $h > 0$ such that the difference quotient on the right of (17) is within an arbitrary $\delta > 0$ of $z_u(u_0, 0)$. Then by (17) we can find an arbitrary small $\epsilon > 0$ so that $z_u(u_0 + \theta h, \epsilon)$ is within δ of this difference quotient. Then $(u_0 + \theta h, \epsilon)$ is a point arbitrarily close to $(u_0, 0)$, while $z_u(u_0 + \theta h, \epsilon)$ is arbitrarily close to (within 2δ of) $z_u(u_0, 0)$. This implies the existence of the desired sequence obeying (30).

If we combine all the preceding reasoning, it results that every sequence (u_n, v_n) must be in case (b), which expresses the required continuity property (20) of z_u.

4. Let $z(u, v)$ be prolonged into the lower half-plane H' $(v < 0)$ by the formula

$$z(u, -v) = z(u, v). \tag{32}$$

Then $z(u, v)$ is defined on the set $H + ab + H'$, i.e., in the complete plane slit along the u-axis from a (inclusive) to $-\infty$ and from b (inclusive) to $+\infty$. We shall denote by S the open region formed by this slit plane: $S = H + ab + H'$.

It is evident that the values (15), (16) of $z_v(u_0, 0)$, $z_u(u_0, 0)$, which apply to z as defined on $H + ab$, are valid also for z as defined on S. It is therefore plain that $z(u, v)$ is of class C' in S, that is, continuous with its first partial derivatives z_u, z_v.

We now apply the following criterion of Bôcher and Koebe[10] to prove that $z(u, v)$ *is harmonic in* S.

CRITERION. *A necessary and sufficient condition that a function* $z(u, v)$ *of class* C' *in an open region* S *be harmonic in* S *is that*

$$\int_K \frac{\partial z}{\partial n}\, ds = 0 \tag{33}$$

for every circumference K *contained in* S.

It is easy to show that the condition (33) (where $\partial/\partial n$ denotes differentiation in the direction of the normal to K) is obeyed by our function $z(u, v)$. Evidently, we have to consider only circles K which traverse the axis-segment ab.

[10] See O. D. Kellogg, *Foundations of Potential Theory*, 1929, p. 227.

Let AB, A_1B_1 be the chords of any such circle which are intercepted from the lines $v = \pm\epsilon$, respectively. These divide the circular disc bounded by K into an upper segment σ, a lower segment σ', and a curvilinear quadrilateral $q(\epsilon) = ABB_1A_1$. We shall denote by the same symbols the boundaries of these regions.

Then

$$\int_K \frac{\partial z}{\partial n}\,ds = \int_\sigma + \int_{\sigma'} + \int_{q(\epsilon)}.$$

In view of the harmonic character of $z(u, v)$ in H, containing σ, and in H', containing σ', the first two terms on the right vanish by Green's theorem, so that

$$\int_K \frac{\partial z}{\partial n}\,ds = \int_{q(\epsilon)} \frac{\partial z}{\partial n}\,ds. \tag{34}$$

We write

$$\int_{q(\epsilon)} \frac{\partial z}{\partial n}\,ds = \int_{AB} + \int_{BB_1} + \int_{B_1A_1} + \int_{A_1A},$$

and allow ϵ to tend to 0. Then each term on the right tends to zero. For AB and B_1A_1, this follows from (14)[11] and the bounded lengths of AB, B_1A_1. For BB_1 and A_1A, we observe that their lengths tend to zero while the integrand stays bounded.

Accordingly, as $\epsilon \to 0$,

$$\int_{q(\epsilon)} \frac{\partial z}{\partial n}\,ds \to 0. \tag{35}$$

The condition (33) follows from (34), holding for every sufficiently small ϵ, and (35).

5. In summary, under the stated hypotheses (i)–(iv) of §2, the formulas (6) and (32), i.e.,

$$x(u, -v) = -x(u, v), \qquad y(u, -v) = -y(u, v), \qquad z(u, -v) = z(u, v), \tag{36}$$

define an analytic prolongation of the harmonic functions $x(u, v)$, $y(u, v)$, $z(u, v)$ across ab. That the other conditions (2), (3) for a minimal surface are satisfied by these prolongations is obvious, since, by differentiation of (36),

$$\begin{aligned} &x_u(u, -v) = -x_u(u, v), \quad y_u(u, -v) = -y_u(u, v), \quad z_u(u, -v) = z_u(u, v), \\ &x_v(u, -v) = x_v(u, v), \qquad y_v(u, -v) = y_v(u, v), \qquad z_v(u, -v) = -z_v(u, v). \end{aligned} \tag{37}$$

[11] The approach there indicated is, of course, uniform on every inner partial segment $a'b'$ of ab; this follows from the continuity of z_v.

The formulas (36) therefore define the required analytic prolongation of the minimal surface M.[12]

6. **An application.** Some years ago the author proved the following theorem:[13] *If any two contours interlace, they are the boundaries of a doubly-connected minimal surface M.*

The result of the present paper may be used to give the following simple construction of M in the case where the given contours are two interlacing circles in an appropriate relative position.

Let any two distinct planes π_1 and π_2 intersect in a line L, on which choose any four points A_1, A_2, B_1, B_2 located in that order. In π_1 construct a semicircle $A_1C_1B_1$ on A_1B_1 as diameter, in π_2 a semicircle $A_2C_2B_2$ on A_2B_2 as diameter. Then form the contour $\Gamma = A_1A_2C_2B_2B_1C_1A_1$, consisting of the two semicircles and the two line segments A_1A_2, B_1B_2. By the one-contour theory,[14] Γ bounds a simply-connected minimal surface M'. Let M' be rotated about the line L through 180°, to give M''. Then $M = M' + M''$ is a doubly-connected minimal surface, which is bounded by the two interlacing circles that result by completion of the original two semicircles. Along the line segments A_1A_2, B_1B_2 the surface M presents no sharp edge or any other type of singularity, but rather a perfect smoothness and regularity. For the portions M', M'' of M which are on opposite sides of A_1A_2 and of B_1B_2 are analytic prolongations of one another, as has been proved in this paper.[15]

7. The general problem, of which the one solved in the present paper is a special case, concerns the analytic prolongation of a minimal surface M contiguous to a given analytic[16] arc C across this analytic arc. That is, in the condition (iv) of §2, the straight line l is replaced by a general analytic arc C, and we inquire then as to the possibility of prolonging the harmonic functions $x(u, v)$, $y(u, v)$, $z(u, v)$ across ab. In the case where C is a plane analytic curve ($z = 0$) and M a contiguous region of this plane, given in conformal representation on the half-plane H (according to (1), (2), (3), where $z(u, v) = 0$), the

[12] The proof just completed can be considerably shortened as follows. From (14) it follows that the harmonic function z_v can be prolonged according to the formula $z_v(u, -v) = -z_v(u, v)$. This implies, by partial integration as to v, the analytic prolongation of z according to the third formula (36). (Suggestion of Professor Marston Morse, in whose seminar this paper was presented by the author.) However, the auxiliary facts brought out by the procedure actually followed here seem of considerable interest for themselves.

[13] J. Douglas, *The problem of Plateau for two contours*, Journ. Math. Phys., Mass. Inst. of Tech., vol. 10(1931), p. 351.

[14] J. Douglas, *Solution of the problem of Plateau*, Trans. Amer. Math. Soc., vol. 33(1931), pp. 263–321; T. Radó, Math. Zeitschrift, vol. 32(1930), pp. 762–796.

[15] Evidently, the reasoning of this article gives two doubly-connected minimal surfaces bounded by the given linking circles; one in each pair of opposite dihedral angles formed by the intersecting planes.

[16] Or one might assume some other appropriate degree of regularity.

analytic prolongation of $x(u, v)$, $y(u, v)$ is always possible, being accomplished by Schwarzian reflection in C. This would seem to lend some plausibility to the possibility of the analytic prolongation of M (and of its conformal representation on H) in the three- (or n-) dimensional case. However, as far as we know, no investigation of this interesting question has so far been published.

INSTITUTE FOR ADVANCED STUDY.

SURFACES IN FOUR-SPACE OF CONSTANT CURVATURE

By Nathaniel Coburn

1. **Introduction.** We shall divide the work into two parts: (A) ruled surfaces V_2 imbedded in four-space of constant curvature S_4; (B) surfaces V_2 imbedded in four-space of constant curvature S_4 and such that the normal curvature locus of V_2 is linear.[1]

In (A) we classify the ruled V_2 in S_4 by means of the normal curvature locus. Two possible cases exist: (1) the normal curvature locus consists of axial points; (2) this locus consists of planar points. If the locus is axial, then by Struik's extension of Segre's theorem,[2] these V_2 are either ruled V_2 in S_3 or developable V_2 in S_4. If the locus is planar, then we show a one-to-one correspondence exists between any ruled V_2 in S_3 and a set of ruled V_2 in S_4, where S_3 and S_4 both have the same curvature K.[3]

In (B) we shall discuss a class of surfaces V_2 in an S_4 of constant curvature K which can be placed into a one-to-one isometric correspondence with any V_2 in an S_3 of constant curvature $K + L^2$.

Finally, we shall show that correspondence theorems of the type mentioned here furnish us with a method of giving existence proofs.

2. **Notation.** In an S_4 we introduce the coördinate system

$$y^\kappa \qquad (\kappa, \lambda, \mu = 1, 2, 3, 4). \tag{2.1}$$

By means of the equations

$$y^\kappa = y^\kappa(u^a) \qquad (a, b, c = 1, 2) \tag{2.2}$$

containing the two essential parameters u^1, u^2, we introduce a two-dimensional manifold in S_4. If the tangent two-dimensional planes E_2 of the surface do not cut the null cone of S_4 in more than a finite number of lines at any point of the surface, then a Riemannian metric is induced in the surface and it can be called a V_2. This last means that we assume the rank of the first fundamental tensor a'_{cb} of the V_2 is two. On the V_2, we introduce two orthogonal non-isotropic

Received March 21, 1938, and July 1, 1938; in revised form, December 3, 1938.

[1] Schouten and Struik, *Einführung in die neueren Methoden der Differentialgeometrie*, Batavia, vol. II, 1938, p. 108. We shall refer to this volume as II.

[2] II, p. 99.

[3] I am deeply indebted to Professor D. J. Struik for his aid in revising part A of this paper.

congruences with unit tangent vector fields

$$\underset{c}{i^\kappa} = \underset{c}{\epsilon}\, \overset{c}{i^\kappa} \qquad \text{(not summed on } c;\ c, b = 1, 2), \tag{2.3}$$

$$\underset{c}{\epsilon} = \pm 1.$$

The sign of $\underset{c}{\epsilon}$ is positive if $\underset{c}{i^\kappa}$ lies in the + domain of a'_{cb} ; negative if $\underset{c}{i^\kappa}$ lies in the − domain. At each point P of the V_2 , there exists a normal plane E_2 , in which we introduce the two orthogonal non-isotropic unit vectors

$$\underset{p}{i^\kappa} = \underset{p}{\epsilon}\, \overset{p}{i^\kappa} \qquad (p, q = 3, 4), \tag{2.4}$$

$$\underset{p}{\epsilon} = \pm 1.$$

We denote the first fundamental tensor of S_4 by

$$a_{\lambda\mu}. \tag{2.5}$$

If the connecting affinor of the V_2 with respect to S_4 is given by

$$B^\kappa_b = \frac{\partial y^\kappa}{\partial u^b} = \partial_b y^\kappa, \quad \partial_b = \frac{\partial}{\partial u^b}, \quad B^{\lambda\kappa}_{bc} = B^\lambda_b B^\kappa_c, \tag{2.6}$$

then

$$a'_{bc} = B^{\lambda\kappa}_{bc} a_{\lambda\kappa}, \tag{2.7}$$

$$B^\kappa_\lambda = \underset{c}{i^\kappa} \underset{c}{i_\lambda}. \tag{2.8}$$

The second fundamental tensors $\underset{p}{h_{cb}}$ of the V_2 are related to the relative curvature tensor $H^{\cdot\cdot\kappa}_{bc}$ of the V_2 with respect to the S_4 by the equations

$$H^{\cdot\cdot\kappa}_{bc} = \underset{p}{h_{bc}}\, \underset{p}{i^\kappa} = \underset{p}{h_{bc}}\, \overset{p}{i^\kappa}, \tag{2.9}$$

where

$$\underset{p}{h_{bc}} = \underset{p}{\epsilon}\, \underset{p}{h_{bc}} = -B^{\lambda\kappa}_{bc} \nabla_\lambda \underset{p}{i_\kappa}, \tag{2.10}$$

∇_λ denoting the covariant differentiation operator of S_4 $(a_{\lambda\mu})$. If the Riemann-Christoffel affinor of the V_2 is K'_{dcba} , and K is the curvature of S_4 , then the Gauss, Codazzi, Ricci relations[4] are

$$K'_{dcba} = \left(\underset{p}{h_{db}}\, \underset{p}{h_{ca}} - \underset{p}{h_{da}}\, \underset{p}{h_{cb}}\right) + K(a'_{db} a'_{ca} - a'_{da} a'_{cb}), \tag{2.11}$$

$$\nabla'_{[c}\, \underset{3}{h_{b]a}} = -v_{[c}\, \underset{4}{h_{b]a}}, \qquad \nabla'_{[c}\, \underset{4}{h_{b]a}} = v_{[c}\, \underset{3}{h_{b]a}}, \tag{2.12}$$

$$\nabla'_{[c}\, v_{b]} = \underset{3}{h^a_{\cdot[c}}\, \underset{4}{h_{b]a}}, \tag{2.13}$$

[4] II, pp. 130, 131.

where ∇'_c is the covariant differentiation operator with respect to a'_{bc} and v_b is a vector field of V_2 given by

$$v_b = B^a_b \left(\nabla_a \underset{3}{i^x}\right) \overset{4}{i_x} = -B^a_b \left(\nabla_a \underset{4}{i^x}\right) \overset{3}{i_x}. \tag{2.14}$$

It is to be noted that these last equations are valid only if $\underset{3}{i^x}$, $\underset{4}{i^x}$ lie in the same domain of directions. Otherwise signs must be changed.

A. Ruled V_2 in S_4

3. **Fundamental relations.** If V_2 is a ruled surface, then it contains a congruence of geodesics of S_4. Let $\underset{1}{i^x}$ be the unit vector field tangent to the geodesics of this congruence. Thus

$$\underset{p}{h_{cb}} \underset{1}{i^b} = \underset{p}{a}\, \underset{2}{i_c}, \tag{3.1}$$

$$\underset{1}{i^b} \nabla'_b \underset{1}{i^c} = 0, \tag{3.2}$$

where $\underset{p}{a}$ are scalar coefficients. The first equation expresses the fact that the rulings have normal curvature equal to zero; the second equation states that the rulings are geodesics of the V_2. As an immediate consequence of (3.1), it follows that the tensors $\underset{p}{h_{cb}}$ can be expressed as follows:

$$\underset{p}{h_{cb}} = \underset{p}{a} \left(\underset{1}{i_c} \underset{2}{i_b} + \underset{2}{i_c} \underset{1}{i_b}\right) + \underset{p}{b}\, \underset{2}{i_c} \underset{2}{i_b}, \tag{3.3}$$

where $\underset{p}{b}$ are scalar coefficients. From (3.3), we see that if the tensor H_{cb} of V_2,

$$H_{cb} = \underset{3}{a^{-1}} \underset{3}{h_{cb}} - \underset{4}{a^{-1}} \underset{4}{h_{cb}}, \tag{3.4}$$

is different from zero, it is of rank 1. Hence we have two cases:

$$H_{cb} = 0, \tag{3.5}$$

$$H_{cb} \underset{1}{i^c} = 0. \tag{3.6}$$

Let us study (3.5). In this case, the equations (3.4) state that the tensor $\underset{3}{h_{cb}}$ is a multiple of $\underset{4}{h_{cb}}$. Hence we have axial points on V_2. From Struik's extension of Segre's theorem[5] *the ruled V_2 in S_4 of the first type have axial points and hence either lie in S_3 or are developable V_2.* Let us now study case (3.6), the case of planar points. Here, there exists a second fundamental tensor of V_2 which is proportional to H_{cb}. To construct this tensor, we introduce into the normal

[5] II, p. 95.

E_2 the orthogonal vectors

$$\overset{3}{i}^{\alpha} = \rho\left(a^{-1}\overset{3}{i}^{\alpha} - a^{-1}\overset{4}{i}^{\alpha}\right), \tag{3.7}$$

$$\overset{4}{i}^{\alpha} = \rho\left(a^{-1}\overset{3}{i}^{\alpha} + a^{-1}\overset{4}{i}^{\alpha}\right), \tag{3.8}$$

where ρ is defined so as to make these unit vectors. Then from (2.10) we see that the corresponding second fundamental tensors $\overset{3}{h}_{cb}$, $\overset{4}{h}_{cb}$ are given by

$$\overset{3}{h}_{cb} = \rho H_{cb}, \tag{3.9}$$

$$\overset{4}{h}_{cb} = \rho\left(a^{-1}\overset{4}{h}_{cb} + a^{-1}\overset{3}{h}_{cb}\right). \tag{3.10}$$

We shall now drop the bar on the h's but use this set of second fundamental tensors in the remainder of the paper. Thus

$$\overset{3}{h}_{cb}\overset{1}{i}^{b} = 0, \tag{3.11}$$

$$\overset{4}{h}_{cb}\overset{1}{i}^{b} = \mu\overset{2}{i}_{c}, \tag{3.12}$$

where μ is a scalar coefficient. Hence

$$\overset{3}{h}_{cb} = \lambda\overset{2}{i}_{c}\overset{2}{i}_{b}, \tag{3.13}$$

$$\overset{4}{h}_{cb} = \mu\left(\overset{1}{i}_{c}\overset{2}{i}_{b} + \overset{2}{i}_{c}\overset{1}{i}_{b}\right) + \nu\overset{2}{i}_{c}\overset{2}{i}_{b}, \tag{3.14}$$

where λ, ν are additional scalar coefficients. From (3.2) it follows that

$$\nabla'_{b}\overset{1}{i}^{c} = -\epsilon\overset{2}{\mathcal{K}}\overset{2}{i}_{b}\overset{2}{i}^{c}, \tag{3.15}$$

where $\overset{2}{\mathcal{K}}$ is the geodesic curvature of the congruence $\overset{2}{i}^{\alpha}$. Covariant differentiation of (3.11) gives

$$\left(\nabla'_{c}\overset{3}{h}_{ba}\right)\overset{1}{i}^{a} + \overset{3}{h}_{ba}\nabla'_{c}\overset{1}{i}^{a} = 0. \tag{3.16}$$

By use of (3.15), the equations (3.16) become

$$\left(\nabla'_{c}\overset{3}{h}_{ba}\right)\overset{1}{i}^{a} - \epsilon\overset{2}{\mathcal{K}}\overset{3}{h}_{ba}\overset{2}{i}_{c}\overset{2}{i}^{a} = 0. \tag{3.17}$$

Alternating on c, b and using the Codazzi relations (2.12), we find that the set (3.17) becomes

$$v_{[c}\overset{4}{h}_{b]a}\overset{1}{i}^{a} = 0. \tag{3.18}$$

By (3.12), these equations become

$$\mu v_{[c}\overset{2}{i}_{a]} = 0. \tag{3.19}$$

This last equation has two solutions

$$\mu = 0, \tag{3.20}$$

$$v^c = k\,\underset{2}{i}{}^c, \tag{3.21}$$

where k is a scalar coefficient. For the case (3.20) from (3.13), (3.14), we see that $\underset{3}{h}_{cb}$ is a multiple of $\underset{4}{h}_{cb}$. This is the case of axial points. Equations (3.21) lead us to ruled V_2 with planar points.

4. **A one-to-one correspondence.** In this section, we shall be concerned with ruled V_2 possessing planar points. By substituting (3.21), (3.13) and (3.14) into the Gauss, Codazzi, Ricci relations (2.11), (2.12), (2.13), the latter become

$$K'_{dcba} = \left(\underset{4}{h}_{db}\underset{4}{h}_{ca} - \underset{4}{h}_{da}\underset{4}{h}_{cb}\right) + K(a'_{db}a'_{ca} - a'_{da}a'_{cb}), \tag{4.1}$$

$$\nabla'_{[c}\,\underset{3}{h}_{b]a} = -k\,\mu\,\underset{2}{i}_{[c}\,\underset{1}{i}_{b]}\,\underset{2}{i}_{a}, \tag{4.2}$$

$$\nabla'_{[c}\,\underset{4}{h}_{b]a} = 0, \tag{4.3}$$

$$\nabla'_{[c}\,k\underset{2}{i}_{b]} = \lambda\mu\,\underset{1}{i}_{[c}\,\underset{2}{i}_{b]}. \tag{4.4}$$

Also from (3.2) and (3.12), we find

$$\underset{4}{h}_{ba}\underset{1}{i}{}^a = \mu\,\underset{2}{i}_b, \qquad \underset{1}{i}{}^b\nabla'_b\,\underset{1}{i}{}^c = 0. \tag{4.5}$$

Now (4.1), (4.3) are the Gauss, Codazzi relations for a V_2 in S_3 of same curvature K as S_4 and with second fundamental tensor $\underset{4}{h}_{cb}$. Hence these conditions and (4.5) determine a ruled V_2 in S_3. If we can show that the equations (4.2), (4.4) have solutions for arbitrary $\underset{4}{h}_{cb}$ (that is, μ, ν), then these last equations will not condition $\underset{4}{h}_{cb}$ and we have the following theorem:

There exists a one-to-one correspondence between any ruled V_2 in S_3 and a special set of isometric ruled V_2 in S_4 of same curvature as S_3 in such a way that the second fundamental tensor of V_2 in S_3 is identical with one of the second fundamental tensors of V_2 in S_4.

We now show that (4.2), (4.4) have solutions for arbitrary μ, ν. The equations (4.2) can be written as

$$\frac{d\lambda}{\underset{1}{ds}} - \underset{2}{\mathcal{K}}\lambda + k\mu = 0 \tag{4.6}$$

by use of (3.13) and (3.14). Likewise (4.4) can be written as

$$\frac{dk}{\underset{1}{ds}} - \underset{2}{\mathcal{K}}k - \lambda\mu = 0. \tag{4.7}$$

By choosing parameter lines along the geodesic $\overset{1}{i}{}^{\alpha}$ and along $\overset{2}{i}{}^{\alpha}$, (4.6) and (4.7) become a system of simultaneous linear differential equations with independent variable u^1 and dependent variables k, λ. Such a system has solutions for all μ. Note that ν does not enter these equations. Hence our theorem is proved.

This theorem shows that the metric properties of ruled V_2 in S_3 and ruled V_2 in S_4 are identical when S_3 and S_4 have the same curvature K.

5. **The curvature vector of the ruled V_2 in S_4.** Consider a unit tangent vector field on V_2 which is determined by the equation

$$i^{\alpha} = \overset{1}{i}\overset{1}{i}{}^{\alpha} + \overset{2}{i}\overset{2}{i}{}^{\alpha} = \overset{b}{i}\overset{b}{i}{}^{\alpha}. \tag{5.1}$$

If the directions i^{α} and $\overset{b}{i}{}^{\alpha}$ both lie in either the $+$ or the $-$ domain of directions of the fundamental tensor a'_{cb}, then

$$\overset{b}{i} = \cos(i^{\alpha}, \overset{b}{i}{}^{\alpha}). \tag{5.2}$$

To this direction i^{α}, there belongs a normal curvature vector u^{κ} which is perpendicular to the V_2 at this point. This vector is the curvature vector of the geodesic curve in the direction i^{α} on the V_2. We determine this vector by means of the relation

$$u^{\kappa} = \epsilon H_{dc}^{\cdot\cdot\kappa} i^{d} i^{c}, \qquad \epsilon = i^{\alpha} i_{\alpha}. \tag{5.3}$$

From (5.1) this last equation becomes

$$u^{\kappa} = (\epsilon H_{dc}^{\cdot\cdot\kappa} \overset{b}{i}{}^{d} \overset{a}{i}{}^{c}) \overset{b}{i}\overset{a}{i}. \tag{5.4}$$

If a'_{cb} is positive definite,[6] then

$$u^{\kappa} = \tfrac{1}{2}(\underset{1}{u}{}^{\kappa} + \underset{2}{u}{}^{\kappa}) + \tfrac{1}{2}(\underset{1}{u}{}^{\kappa} - \underset{2}{u}{}^{\kappa}) \cos 2\phi + \underset{12}{u}{}^{\kappa} \sin 2\phi, \tag{5.5}$$

where by (2.9), (3.13), (3.14)

$$\overset{1}{i} = \cos\phi, \qquad \overset{2}{i} = \sin\phi, \tag{5.6}$$

$$\underset{1}{u}{}^{\kappa} = H_{bc}^{\cdot\cdot\kappa} \overset{b}{i}\overset{c}{i} = 0, \tag{5.7}$$

$$\underset{2}{u}{}^{\kappa} = H_{bc}^{\cdot\cdot\kappa} \overset{b}{i}\overset{c}{i} = \lambda \underset{3}{i}{}^{\kappa} + \nu \underset{4}{i}{}^{\kappa}, \tag{5.8}$$

$$\underset{12}{u}{}^{\kappa} = H_{bc}^{\cdot\cdot\kappa} \overset{b}{i}\overset{c}{i} = \mu \underset{4}{i}{}^{\kappa}. \tag{5.9}$$

Hence

$$u^{\kappa} = \left(\frac{\lambda}{2} - \frac{\lambda}{2}\cos 2\phi\right) \underset{3}{i}{}^{\kappa} + \left(\frac{\nu}{2} - \frac{\nu}{2}\cos 2\phi + \mu \sin 2\phi\right) \underset{4}{i}{}^{\kappa}. \tag{5.10}$$

[6] II, p. 105.

Let us study the asymptotic directions on the ruled V_2, that is, the directions for which u^α vanishes. We have two possible solutions

$$\phi = 0, \tag{5.11}$$

$$\lambda = 0, \qquad \cot\phi = -\frac{\nu}{2\mu}. \tag{5.12}$$

The case (5.11) leads us to the conclusion that $\underset{1}{i}{}^\alpha$ is a unit tangent vector in an asymptotic direction. In order to determine whether other asymptotic directions exist, we analyze (5.12). From (4.2) and (3.13), we find that the two possible solutions are

$$\mu = 0, \tag{5.13}$$

$$k = 0. \tag{5.14}$$

In case (5.13), from (3.14) we find that both second fundamental tensors are of rank 1. From Segre's theorem, these V_2 are developable. If (5.14) is satisfied, by use of (5.12), (3.13), (3.21) we find

$$\underset{3}{h}_{cb} = 0, \qquad v_c = 0. \tag{5.15}$$

These V_2 lie in S_3.[7]

In the case where a'_{cb} is indefinite, a similar analysis leads to the same results.

Our conclusions can be summarized as follows: *The only case in which a ruled V_2 in S_4 has, apart from its rulings, a second set of asymptotic lines is when the V_2 is imbedded in an S_3. In particular, all doubly ruled V_2 in S_4 are V_2 in S_3.*

B. Surfaces with linear points

6. **V_2 in S_4 with linear non-axial points.**[8] If we let x, y represent the projections of the normal curvature vector u^α on $\underset{3}{i}{}^\alpha$, $\underset{4}{i}{}^\alpha$, then

$$x = u^\alpha \underset{3}{i}_\alpha = \underset{3}{h}_{cb} i^c i^b, \tag{6.1}$$

$$y = u^\alpha \underset{4}{i}_\alpha = \underset{4}{h}_{cb} i^c i^b. \tag{6.2}$$

In the case of axial points, these V_2 are developable or V_2 in S_3. If the locus consists of linear non-axial points, then for *all* i^α (arbitrary unit vectors on V_2)

$$y - mx + b = 0 \qquad (b \neq 0). \tag{6.3}$$

This equation can be written

$$\left(\underset{4}{h}_{cb} - m\underset{3}{h}_{cb} + ba'_{cb}\right) i^c i^b = 0, \tag{6.4}$$

$$\underset{4}{h}_{cb} - m\underset{3}{h}_{cb} + ba'_{cb} = 0. \tag{6.5}$$

[7] II, p. 50.

[8] II, p. 108, A5; p. 109, B5.

When we transform from one set of orthogonal vectors $\underset{p}{i}^{\alpha}$ in the normal E_2 to another set $\underset{p}{i}^{*\alpha}$, we obtain the equations

$$\underset{p}{i}^{*\alpha} = \overset{q}{\underset{p}{\alpha}}\, \underset{q}{i}^{\alpha}. \tag{6.6}$$

From (2.10), we find

$$\underset{p}{\bar{h}}_{ca} = \overset{q}{\underset{p}{\alpha}}\, h_{ca}. \tag{6.7}$$

If we write (6.5) in the form

$$\rho b a'_{cb} = \rho \left(m \underset{3}{h}_{cb} - \underset{4}{h}_{cb} \right), \tag{6.8}$$

where

$$\rho = (1 + m^2)^{-\frac{1}{2}}, \tag{6.9}$$

and define $\overset{q}{\underset{p}{\alpha}}$ by

$$\begin{aligned} \overset{3}{\underset{3}{\alpha}} &= \rho m, & \overset{4}{\underset{3}{\alpha}} &= -\rho, \\ \overset{3}{\underset{4}{\alpha}} &= \rho, & \overset{4}{\underset{4}{\alpha}} &= \rho m, \end{aligned} \tag{6.10}$$

then (6.7) becomes

$$\underset{3}{\bar{h}}_{ca} = \rho b a'_{cb}. \tag{6.11}$$

We shall drop the bar in our future work but use the set of normal vectors $\underset{p}{i}^{\alpha}$ for which (6.11) is valid. Furthermore, we replace ρb by L and write (6.11) as

$$\underset{3}{h}_{cb} = L a'_{cb}. \tag{6.12}$$

From elementary geometry, we easily see that

$$L = \rho b \tag{6.13}$$

is the perpendicular distance from the origin to the line (6.3)—assuming $\underset{3}{i}^{\alpha}$, $\underset{4}{i}^{\alpha}$ both lie in the same domain of directions (+ or −). We shall study those V_2 for which the quantity L is constant over V_2. These V_2 are characterized by the fact that their normal curvature loci at various points P of V_2 consist of the tangents to a circle of constant radius L in the normal E_2.

7. **A correspondence theorem for these V_2.** Substituting (6.12) into (2.12), we obtain

$$-v_{[c}\, \underset{4}{h}_{b]a} = 0, \tag{7.1}$$

since L is constant. But (7.1) implies either

$$v_c = 0, \tag{7.2}$$

or

$$\overset{4}{h}_{ba} \text{ is of rank 1.} \tag{7.3}$$

For the case (7.2), the Gauss, Codazzi, Ricci relations become

$$K'_{dcba} = \left(\overset{4}{h}_{db}\overset{4}{h}_{ca} - \overset{4}{h}_{da}\overset{4}{h}_{cb}\right) + (K + L^2)(a'_{db}a'_{ca} - a'_{da}a'_{cb}), \tag{7.4}$$

$$\nabla'_{[c}\overset{4}{h}_{b]a} = 0. \tag{7.5}$$

But these are the Gauss, Codazzi relations for any V_2 $\left(\overset{4}{h}_{cb} \text{ arbitrary}\right)$ in S_3 of constant curvature $K + L^2$. If (7.3) is valid, then the Gauss equations show that V_2 is of constant curvature $K + L^2$. Hence, we have the following theorem:

There exists a one-to-one correspondence between any V_2 in an S_3 of constant curvature $K + L^2$ and an isometric V_2 in an S_4 of constant curvature K, whose normal curvature loci consist of the ∞^1 tangent lines to a circle of radius L about P in the normal E_2, unless V_2 is an S_2 of constant curvature $K + L^2$. This correspondence is of such nature that the second fundamental tensor of V_2 in S_3 is identical with one of the second fundamental tensors of V_2 in S_4.

Since all our work in §§6 and 7 depended on the properties of the normal E_2, these last results are valid when V_2 is replaced by V_m, S_3 by S_{m+1}, and S_4 by S_{m+2}.

In concluding, we wish to point out that correspondence theorems of the type discussed in this paper furnish us with a method of giving existence proofs. The question arises: "Do V_2 in S_4 of the type discussed in §§6 and 7 exist?" By our work, the existence of these V_2 in S_4 depends on the existence of isometric V_2 in a certain S_3. From Schläfli's[9] theorem into any S_3 any metrical V_2 may be imbedded. Hence, our V_2 in S_4 exist. Furthermore, these V_2 in S_4 are "metrically unlimited"; their metric tensor is arbitrary. Similarly, V_m in S_{m+2} (of curvature K) with the same normal curvature properties exist. However, these V_m are "metrically limited" to those metrics which a specific S_{m+1} (curvature $K + L^2$) will admit.

UNIVERSITY OF TEXAS.

[9] II, p. 142.

THE CHARACTERIZATION OF FLAT PROJECTIVE SPACES BY THEIR DIFFERENTIAL INVARIANTS

By Robert Singleton

Introduction. For the ordinary projective space, Veblen has derived a flat projective connection[1, 2] and certain differential conditions satisfied by this connection. He has shown that these differential conditions are sufficient to characterize locally the ordinary projective space. T. Y. Thomas, in lectures at Princeton University in the fall of 1937, proved that the connection could be defined over a space homeomorphic to the ordinary projective space. In this paper it is shown that the ordinary projective space is completely characterized by the above differential conditions on the projective connection. This is done by finding a set of projective coördinate systems extending to the entire space. Difficulty is encountered in extending the coördinate systems uniquely since the space is not simply connected. The only other point of possible difficulty comes in showing that the coördinate systems assume each arithmetic value only once as they cover the space.

I want to express my gratitude to Professor T. Y. Thomas for his material aid with criticisms and suggestions during the writing of this paper.

1. Consider an n-dimensional space, P_n, which is homeomorphic to the straight lines through the origin in an $(n+1)$-dimensional number space, A_{n+1}. P_n is coördinated with spherical coördinate systems such that the transformations between coördinate systems in the overlapping regions are analytic; and a projective connection Π is defined on P_n with components $\Pi^{\alpha}_{\beta\gamma}$ [Greek indices range from 0 to n, Latin from 1 to n] which are analytic functions of the coordinates x^i of the space, and transform mechanically like an affine connection. That is, for the transformation $x^i = x^i(\bar{x})$,

$$\bar{\Pi}^{\alpha}_{\beta\gamma}\frac{\partial \bar{x}^{\lambda}}{\partial \bar{x}^{\alpha}} = \Pi^{\lambda}_{\mu\nu}\frac{\partial x^{\mu}}{\partial \bar{x}^{\beta}}\frac{\partial x^{\nu}}{\partial \bar{x}^{\gamma}} + \frac{\partial^2 x^{\lambda}}{\partial \bar{x}^{\beta}\,\partial \bar{x}^{\gamma}},$$

where the derivatives which include an x^0 are defined by

$$\frac{\partial x^{\lambda}}{\partial \bar{x}^{0}} = \delta^{\lambda}_{0}, \qquad \frac{\partial x^{0}}{\partial \bar{x}^{i}} = \frac{1}{\rho(\bar{x})}\frac{\partial \rho(\bar{x})}{\partial \bar{x}^{i}},$$

Received May 23, 1938.

[1] O. Veblen, *Projektive Relativitätstheorie*, Ergebnisse der Mathematik, vol. 2, 1933.

[2] O. Veblen, *Generalized projective geometry*, Journal of the London Mathematical Society, vol. 4(1929), p. 140.

$\rho(\bar{x})$ being an arbitrary analytic function[3] of $\bar{x}^1, \bar{x}^2, \cdots, \bar{x}^n$. Furthermore,

$$\Pi^\alpha_{\beta\gamma} = \Pi^\alpha_{\gamma\beta}; \qquad \Pi^\alpha_{\beta 0} = \delta^\alpha_\beta$$

and

$$R^\alpha_{\beta\gamma\delta} = \frac{\partial \Pi^\alpha_{\beta\gamma}}{\partial x^\delta} - \frac{\partial \Pi^\alpha_{\beta\delta}}{\partial x^\gamma} + \Pi^\sigma_{\beta\gamma}\Pi^\alpha_{\sigma\delta} - \Pi^\sigma_{\beta\delta}\Pi^\alpha_{\sigma\gamma} = 0.$$

In this definition

$$\frac{\partial \Pi^\alpha_{\beta\gamma}}{\partial x^0} = 0.$$

It is to be shown that in the space of structure given above, preferred coordinate systems can be defined which satisfy the axioms for a projective space. Specifically, preferred coördinate systems are those which have the following properties:[4]

1. In a preferred coördinate system each point is represented by at least one arithmetic point in a space of dimension $n + 1$, and each arithmetic point other than the origin represents just one point.

2. Two arithmetic points other than the origin represent the same point if and only if they lie on the same arithmetic straight line through the origin.

3. Any preferred coördinate system can be transformed into any other by a linear homogeneous transformation.

4. Any coördinate system obtained from a preferred coördinate system by a linear homogeneous transformation is preferred.

2. The preferred coördinate systems arise from the equations

$$\frac{\partial^2 Z}{\partial x^\alpha \partial x^\beta} = \Pi^\sigma_{\alpha\beta}\frac{\partial Z}{\partial x^\sigma}, \tag{1a}$$

$$\frac{\partial Z}{\partial x^0} = Z. \tag{1b}$$

Equation (1b) shows that solutions are of the special form $Z = e^{x^0} f(x^i)$, and hence the relationship between solutions is linear homogeneous instead of the general linear.

The properties of $\Pi^\alpha_{\beta\gamma}$ and the coördinate systems assure that a solution for equations (1) exists in a sufficiently small neighborhood of any point x^i of P_n as an analytic function of $x^0, \cdots, x^n$, determined uniquely by choosing an x^0 and initial values for $\partial Z/\partial x^\alpha$ [that is, $\partial Z/\partial x^i$ and Z]. Any $n + 1$ such functions

[3] T. Y. Thomas has shown for the flat case that a gauge x^0 may always be selected such that, in each transformation between coördinate neighborhoods, $\rho(x) = (\bar{x}/x)$, the Jacobian of the coördinate transformation. T. Y. Thomas, *Recent Trends in Geometry*, in *Semicentennial Addresses of the American Mathematical Society*, vol. 2, New York, 1938.

[4] Veblen and Whitehead, *Foundations of Differential Geometry*, Cambridge University Press, 1932.

Z^α are independent if the $n+1$ initial values $\partial Z^\alpha/\partial x^\beta$ are independent vectors. Any solution $\overline{Z}$ is related to an independent set Z^α by $\overline{Z} = p_\alpha Z^\alpha$, and any such homogeneous linear combination is a solution. $\overline{Z}^\beta = p^\beta_\alpha Z^\alpha$ are independent if Z^α are independent and $|p| \neq 0$. Hence, independent sets of solutions of equations (1) have properties 3 and 4.

From the form $Z = e^{x^0} f(x^i)$ it is seen that the choice of initial conditions determines f as an analytic function of the coördinates in the neighborhood of the initial point. Thus, we may say that $f(x^i)$ is determined uniquely by the choice of initial values for $\partial f/\partial x^i$ and f, and Z is determined by that choice, except for the arbitrariness permitted by the factor e^{x^0}. Independence of solutions is not disturbed by this change in point of view.

A solution $Z^1(x)$ which exists in a neighborhood $N(x_1)$ about a point x_1 may be shown to exist at every point of P_n, the usual extension of such functions being used as a definition of $Z^1(x)$ at points outside $N(x_1)$. A point $x_2 \subset N(x_1)$ [called a "center of extension"] is selected, and by the theory of the preceding paragraph a function $Z^2(x)$ for which $f^2(x_2) = kf^1(x_2)$ and

$$\frac{\partial f^2(x_2)}{\partial x^i} = k \frac{\partial f^1(x_2)}{\partial x^i}$$

can be found in a neighborhood $N(x_2)$, and hence coinciding with $Z^1(x)$ in the connected part of the intersection of $N(x_1)$ and $N(x_2)$ containing x_2. Z^2 is called an extension of Z^1 to points of $N(x_2)$ outside $N(x_1)$.

There is no point to which Z^1 can not be extended by this method. Let any such points be denoted by y, and the points to which Z^1 can be extended by x. Then because of the method of extension by neighborhoods, $\{x\}$ is open. Hence $\{y\}$ is closed. Hence in every neighborhood of some point y_1 exist points of $\{x\}$. Since every point may serve as an initial point for equations (1), in some $N(y_1)$ there exist $n+1$ independent solutions of (1), $\overline{Z}^\alpha(x^i)$ and also a point x_0 at which Z^1 exists. Then set

$$a_\alpha \frac{\partial \overline{Z}^\alpha(x_0)}{\partial x^i} = \frac{\partial Z^1(x_0)}{\partial x^i}, \qquad a_\alpha \overline{Z}^\alpha(x_0) = Z^1(x_0).$$

Since $\overline{Z}^\alpha$ are independent, these equations may be solved for a_α, and hence $\overline{Z} = a_\alpha \overline{Z}^\alpha$ is an extension of Z^1. Then Z^1 is extended to y_1, since $\overline{Z}$ exists at y_1, contrary to hypothesis. No points y_1 can exist, the space being connected.

We note here that P_n is arcwise connected, compact and bicompact, since it is homeomorphic to the lines through the origin of an $(n+1)$-dimensional arithmetic space.

This extension does not disturb the independence of an originally independent set of solutions. If at a point, $J = |\partial Z^\alpha/\partial x^\beta| \neq 0$, dJ/dt can be integrated along a curve $x^i(t)$ to give

$$J(t) = \frac{1}{n+1} J(t_0) \exp\left[\int_{t_0}^{t} \Pi^\alpha_{\alpha i} \frac{dx^i}{dt}\, dt\right].$$

3. In considering uniqueness of extension we need a notion of two different kinds of sets of neighborhoods. Suppose a function Z at an initial point x_1 to be extended to a point x_2 through two different sets of neighborhoods. It is necessary to prove that the same element at x_2 results from both extensions. Since P_n is not simply connected, a curve without double points joining x_1 and x_2 lying entirely within one set may or may not be deformable into such a curve in the other set of neighborhoods. No more complicated case arises, for it can easily be shown that a curve can be deformed into any other curve lying in the same set of neighborhoods, each neighborhood being spherical.

The problem of uniqueness of extension from x_1 to x_2 is equivalent to the problem of whether or not the functional element of Z in $N(x_1)$ produces the same element in $N(x_1)$ after extension from $N(x_1)$ to $N(x_1)$ through neighborhoods outside $N(x_1)$. The set of neighborhoods is termed "bounding" or "non-bounding" according as every simple closed curve lying in the set of neighborhoods is or is not a bounding cycle.

It is easily seen that extension through a bounding set of neighborhoods reproduces the same element in $N(x_1)$. A closed curve lying on each of the centers of extension may be deformed to lie entirely in $N(x_1)$. A set of curves may be selected from the positions of this deformation, and centers of extension on each curve so selected, in such a way that the set of neighborhoods associated with each curve overlaps the set of the preceding curve, and hence produces the same element as the preceding set. Since the curve eventually lies within $N(x_1)$, the extension must in each case reproduce the original Z.

Where the set of neighborhoods is not bounding, the proof is not casual. We prove that from each point issue curves (straight lines) of a family, and a unique hypersurface (hyperplane) along which f^1/f^2 maintains a constant value in extension, $e^{x^0}f^1$ and $e^{x^0}f^2$ being two independent solutions of equations (1). These curves and surfaces are then shown to be non-bounding by use of an auxiliary curve.

To simplify the work, we show first that gauges and coördinate systems can be found in which $\Pi^0_{ij} = 0$, $\Pi^k_{ij} = 0$ everywhere on P_n. Under the gauge transformation $\bar{x}^0 = x^0 - \log \rho(x^i)$, Π^0_{ij} transforms by the law

$$\bar{\Pi}^0_{ij} = \Pi^0_{ij} - \Pi^k_{ij}\sigma_k - \sigma_i\sigma_j + \sigma_{ij},$$

where

$$\sigma_i = \frac{1}{\rho}\frac{\partial \rho}{\partial x^i}, \qquad \sigma_{ij} = \frac{\partial \sigma_i}{\partial x^j}.$$

If $\bar{\Pi}^0_{ij}$ is to be zero, σ must satisfy

$$\sigma_{ij} = \Pi^k_{ij}\sigma_k + \sigma_i\sigma_j - \Pi^0_{ij}.$$

The integrability conditions for these equations are

$$\frac{\partial \sigma_{ij}}{\partial x^k} - \frac{\partial \sigma_{ik}}{\partial x^j} = \sigma_s R^s_{ijk} - R^0_{ijk} = 0$$

which are satisfied since $R^{\alpha}_{\beta\gamma\delta} = 0$. Thus, σ_i's and hence functions ρ exist everywhere on P_n, such that Π^0_{ij} can be made zero by the appropriate gauge transformation.

Under the coördinate transformation $\bar{x}^i = \bar{x}^i(x)$, Π^k_{ij} transforms as an affine connection and Π^0_{ij} as a quadratic covariant tensor. If $\bar{\Pi}^k_{ij}$ is to be zero, we need functions $\bar{x}^i$ satisfying

$$\frac{\partial^2 y}{\partial x^i \partial x^j} = \Pi^k_{ij} \frac{\partial y}{\partial x^k}$$

for which the integrability conditions are

$$\frac{\partial y_{ij}}{\partial x^k} - \frac{\partial y_{ik}}{\partial x^j} = \frac{\partial y}{\partial x^s} R^s_{ijk} = 0$$

which are satisfied since $R^{\alpha}_{\beta\gamma\delta} = 0$. Hence n independent solutions exist in a neighborhood of every point; if we set $\bar{x}^i = y^i(x)$, then $\Pi^i_{jk} = 0$. Since Π^0_{ij} is a tensor in coördinate transformation, if it was made zero first by selecting a proper gauge, it will remain zero in the new coördinates.

We now use the functions $f(x^i)$ in the solutions $Z = e^{x^0} f(x^i)$ of equations (1). From these equations f is given in each coördinate system and gauge by

$$f_{ij} = \Pi^s_{ij} f_s + \Pi^0_{ij} f, \tag{2}$$

where

$$f_s = \frac{\partial f}{\partial x^s}, \qquad f_{ij} = \frac{\partial^2 f}{\partial x^i \partial x^j}.$$

With the definition (2) f is a scalar under coördinate transformation, but under the gauge transformation $\bar{x}^0 = x^0 + \log \rho(x)$ we have $f = \rho \bar{f}$. Furthermore, along any curve $x^i(t)$, f is given [from equations (2)] by

$$\dot{f} = f_s(\dot{x}^s + \Pi^s_{ij}\dot{x}^i\dot{x}^j) + f\Pi^0_{ij}\dot{x}^i\dot{x}^j,$$

where the dots stand for ordinary derivatives with respect to t; for f transforming by $f = \rho \bar{f}$ this equation is invariant. However, the straight lines will arise from solutions of

$$\ddot{x}^s + \Pi^s_{ij}\dot{x}^i\dot{x}^j = 0, \tag{3}$$

an equation which is invariant under coördinate transformations, but by the gauge transformation $\bar{x}^0 = x^0 + \log \rho(x)$ goes into

$$\ddot{\bar{x}}^s + \Pi^s_{ij}\dot{\bar{x}}^i\dot{\bar{x}}^j = -\frac{2}{\rho}\dot{\rho}\dot{\bar{x}}^s.$$

Hence when we define curves by (3) in a particular gauge, we will understand that after transformation to another gauge the curve will be defined as just given.

Suppose, then, that gauges and coördinate systems have been selected so that $\Pi^0_{ij} = 0$, $\Pi^k_{ij} = 0$ everywhere on P_n, and consider in any coördinate neighborhood

the curves defined by $\ddot{x}^i = 0$ [equation (3)]. A curve of this family is determined by an initial point x_0^i and an initial direction ξ^i; then $x^i = \xi^i t + x_0^i$. Also $\dot{f} = f_{,s}\dot{x}^s = 0$ along the curve, and $f = f_{,i}(x_0)\xi^i t + f(x_0) = at + b$.

Now suppose we have two independent solutions f^1 and f^2 of (2), whose initial conditions are given at x_0 with $f^2 \neq 0$ at x_0, and consider the function $\phi(x) = f^1/f^2$; ϕ is invariant under all transformations which are admissible. Along the curve $x^i = \xi^i t + x_0^i$,

$$\phi = \frac{a^1 t + b^1}{a^2 t + b^2}; \qquad \dot{\phi} = \frac{a^1 b^2 - a^2 b^1}{(a^2 t + b^2)^2}. \tag{4}$$

ϕ is constant if $\dot{\phi} = 0$; i.e., if $a^1 = kb^1$, $a^2 = kb^2$. Thus, the initial directions for curves issuing from x_0 along which ϕ is constant are the solutions ξ_a^i of

$$\begin{aligned} \xi^i f^1_{,i}(x_0) &= k f^1(x_0), \\ \xi^i f^2_{,i}(x_0) &= k f^2(x_0). \end{aligned} \tag{5}$$

Since some determinant

$$\begin{vmatrix} f^1 & f^1_{,i} \\ f^2 & f^2_{,i} \end{vmatrix} \quad \text{or} \quad \begin{vmatrix} f^1_{,i} & f^1_{,j} \\ f^2_{,i} & f^2_{,j} \end{vmatrix}$$

is not zero, f^1 and f^2 being independent, there exist $n - 1$ linearly independent homogeneous solutions ξ_a^i, k_a $(a = 1, \cdots, n-1)$, and $\lambda^a \xi_a^i$, $\lambda^a k_a$ is also a solution for any $n - 1$ constants λ^a. Thus along all curves

$$x^i = \lambda^a \xi_a^i t + x_0^i \tag{6}$$

ϕ is constant. If $\lambda^a = p\bar{\lambda}^a$ and $\bar{t} = pt$, then

$$x^i = \lambda^a \xi_a^i t + x_0^i = \bar{\lambda}^a \xi_a^i \bar{t} + x_0^i.$$

Hence all solutions proportional to one solution determine the same curve, with different parametrization.

Fixing t at a constant value sufficiently small so that the points of (6) for that value remain in the coördinate neighborhood, and absorbing it into the λ^a, we have from (6)

$$x^i = \lambda^a \xi_a^i + x_0^i, \tag{7}$$

and letting λ^a take all values, we have the same set of points as is given in (6). Thus, (7) are the equations of the $(n-1)$-dimensional hypersurface on which ϕ is constant. It is $(n-1)$-dimensional everywhere since $\| \partial x^i/\partial \lambda^a \| = \| \xi_a^i \|$ is of rank $n - 1$ for any point λ^a. Also, at any point the surface satisfies

$$\frac{\partial \phi}{\partial \lambda^a} = \frac{\partial \phi}{\partial x^i} \cdot \frac{\partial x^i}{\partial \lambda^a} = 0.$$

However, if η^i is any vector such that $\eta^i \neq p^a\xi^i_a$ for any set of constants p^a, then $(\partial\phi/\partial x^i)\eta^i \neq 0$ at any point of the surface. Otherwise for some point

$$\frac{\partial\phi}{\partial x^i}\eta^i = 0, \qquad \frac{\partial\phi}{\partial x^i}\xi^i_a = 0$$

and hence $\partial\phi/\partial x^i = 0$. From this, $f^2f^1_i - f^1f^2_i = 0$, and $f^1_i = k_if^1$, $f^2_i = k_if^2$, which is impossible if f^1 and f^2 are independent. But $(\partial\phi/\partial x^i)\eta^i$ is the derivative of ϕ in a direction η^i away from the surface. Hence for a sufficiently small open set in P_n about any point x^i of the surface, the surface passes through *all* points at which ϕ has the value $\phi(x^i)$.

If the initial point x^i_0 is sufficiently close to points at which $f^1 = f^2 = 0$, on each one of the curves (6) the point for which $t = -b^1/a^1 = -b^2/a^2 = -1/k$ is such a point. Furthermore, any point x^i_1 at which $f^1 = f^2 = 0$ also lies on one of the curves, if it is sufficiently close to x_0. A direction η^i can be found for which the curve $x^i(t) = \eta^it + x^i_0$ passes through x_1, and along this curve also, $f^1 = a^1t + b^1$, $f^2 = a^2t + b^2$, for some a^1, a^2, b^1, b^2. Since $x^i(t_1) = x^i_1$ for some value t_1, and since at x^i_1, $f^1 = a^1t_1 + b^1 = 0$, and $f^2 = a^2t_1 + b^2 = 0$, we have that $t_1 = -b^1/a^1 = -b^2/a^2$. Hence $a^1b^2 - a^2b^1 = 0$, and from (4), ϕ is constant along $x^i(t)$. Thus x^i_1 lies on the surface (7) generated by curves (6), and so the surface issuing from x_0 passes through all such points for a sufficiently restricted neighborhood.

Let now x^i_0 be a point at which $f^1 = f^2 = 0$. At x^i_0, ϕ is indeterminate, and a value may be assigned depending on the path of approach. For one of the curves which are solutions of $\ddot{x}^i = 0$, from (4), $\phi = a^1/a^2 = \xi^if^1_i(x_0)/\xi^if^2_i(x_0)$, and if it is desired that ϕ have the value m/n, then as before there are $n - 1$ independent solutions of

$$\begin{aligned} \xi^if^1_i(x_0) &= mk, \\ \xi^if^2_i(x_0) &= nk \end{aligned} \tag{8}$$

similar to (5), and along the curves $x^i = \xi^it + x^i_0$, ϕ has the value m/n. There are, however, $n - 2$ independent directions for which $\xi^if^1_i(x_0) = \xi^if^2_i(x_0) = 0$, and these lie in the $(n - 2)$-subspace which is the intersection of the surfaces $f^1 = 0$ and $f^2 = 0$. Furthermore, the surface $x^i = \lambda^a\xi^i_a + x^i_0$, determined by $n - 1$ independent solutions ξ^i_a, k_a of (8) also passes through this intersection. For there are $n - 2$ sets of constants $\bar{\lambda}^a$ satisfying $\bar{\lambda}^ak_a = 0$, and for these, $\bar{\lambda}^a\xi^i_af^1_i(x_0) = 0$, $\bar{\lambda}^a\xi^i_af^2_i(x_0) = 0$. Thus, the curves in the directions $\bar{\lambda}^a\xi^i_a$ lie in the subspace $f^1 = f^2 = 0$. We have that at the point x^i_0, where $f^1 = f^2 = 0$, any direction ξ^i gives a curve $x^i = \xi^it + x^i_0$ along which ϕ is constant or indeterminate; and that all surfaces arising from such curves in the manner of the surface (7) [whether or not the generating curves issue from x^i_0] pass through x^i_0 and the intersection of the surfaces $f^1 = 0, f^2 = 0$ if they pass through points sufficiently close to x^i_0.

From the above it is seen that we can impose a two-point boundary condition

on the solutions of $\ddot{x}^i = 0$, to find the solution [and hence also the surface of form (7)] issuing from x_0 [where $f^1 = f^2 = 0$], and passing through any other sufficiently nearby point. Since the solutions are analytic functions of parameters giving the initial direction, such a point is in an open set of points of P_n, to each of which solutions can be found issuing from x_0. The solutions, and hence the surfaces, are also analytic functions of parameters giving the initial point.

Now let us take any non-bounding linear cycle $y^i(t)$ for which $y^i(0) = y^i(1) = x_0^i$. Then by the above development, a surface of the linear pencil generated by $f^1 = 0, f^2 = 0$ can be found including both x_0^i and any nearby point of the curve $y^i(t)$, possibly only up to a point y_1^i, at which we suppose not both f^1 and f^2 are zero. Let ϕ have the value ϕ_1 at y_1. Since y_1 is a boundary point of the linear open set of points y^i through which surfaces can pass from x_0, there exist such points in every neighborhood of y_1. Let such a point sufficiently near to y_1 be y_2.

There exists a surface S_1 issuing from y_1 along which ϕ has the constant value ϕ_1, and by our assumption S_1 does not contain x_0. Since a surface coming sufficiently close to x_0 passes through x_0, there exist closed sets $\bar{N}$ containing S_1 but not containing x_0^i. However, since surfaces generated by the family of solutions of $\ddot{x}^i = 0$, some of which generate S_1, are also given by analytic functions of parameters $\mu, \nu, \cdots$ determining the initial point, if y_2 is sufficiently close to y_1, values of $\mu, \nu, \cdots$ can be found for which y_2 is the initial point, giving rise to the surface S_2; and open sets M containing S_1 can be found which also contain S_2 and are contained in $\bar{N}$ for any desired regions of S_1 and S_2. But there is only one such surface through y_2 and this passes through x_0, y_2 being a point of that character. From this we have the contradiction $\bar{N} \not\supset x_0$ and $\bar{N} \supset M \supset x_0$. Hence no such point y_1 can exist, unless at that point $f^1 = f^2 = 0$. But if this happens, y_1 possesses the same properties as x_0 and so a surface can be found issuing from y_1 containing y_2 and hence also x_0.

Thus, a surface along which ϕ is constant can be found issuing from x_0 and containing any point of the curve $y(t)$, in such a manner that two surfaces lie as close together as desired if the points of the curve to which the two were issued are sufficiently close. Let us designate the surface containing the point $y(m)$ by S_m, and select a curve C_m lying in S_m joining x_0 to $y(m)$. It is seen that C_m and the portion of $y(t)$ between $y(0)$ and $y(m)$ form a bounding cycle, since it can be shrunk to x_0 by reducing m continuously. Then for $t = 1$, $y(1) = x_0$, and so C_1 is a cycle, but is not merely the point x_0, since the curve $y(t)$ is not bounding but $y(t)$ plus C_1 is bounding. For this reason also, C_1 is not bounding.

Now ϕ remains constant as f^1 and f^2 are extended to points of C_1. Hence, in extension from any point of C_1 back to the same point, through sets of neighborhoods which contain C_1, $f^1 \to kf^1$, $f^2 \to kf^2$. Since $y(t)$ could be any non-bounding cycle, the extension through any set of non-bounding neighborhoods produces the same result. Since on a projective manifold twice any non-bounding linear cycle is bounding, $k^2 = 1$.

This can be made to apply to any pair of the functions f^m; hence, in extension

through any set of non-bounding neighborhoods, the f^α are changed at most in sign, and if so, simultaneously. This will not affect the desired homeomorphism between points in P_n and lines in A_{n+1}.

4. The properties 1 and 2 in §1 may be replaced by the equivalent properties:

1a. Each point is represented by the arithmetic points of one line through the origin in an arithmetic space of $n + 1$ dimensions, in each case excluding the origin.

2a. Each line through the origin in an arithmetic space of $n + 1$ dimensions represents just one point.

Thus, the space which we are using is the ordinary arithmetic space, but with the zero point removed. Let this be called B_{n+1}.

We have represented the points of the manifold by x^i. Represent by Z^α whatever arithmetic points are denoted by the values $Z^\alpha = \pm e^{z^0} f^\alpha(x^i)$, where $e^{z^0} f^\alpha(x^i)$ is a particular set of $n + 1$ independent solutions of (1), as x^i covers P_n and $-\infty < x^0 < +\infty$. The set Z^α has already been shown to have the property 1a, because of the form of solution, $e^{z^0} f(x^i)$, since it is defined uniquely everywhere on P_n and since not all $f^\alpha = 0$ at once. To show property 2a it is sufficient to show that the values which Z^α take on fill B_{n+1}, and that no two points x have the same or proportional Z coördinates.

For ease in discussion we call any pair of lines in B_{n+1}, which by the addition of the origin form one straight line in A_{n+1} passing through the origin, a *central line*, and a *line neighborhood* will mean the set of points of an open set of central lines.

First, since Z^α are analytic functions of the x's, and the Jacobian is everywhere different from zero, there is a one-to-one correspondence between the central lines passing through the points of a neighborhood of any point Z^α and the points of a neighborhood of a corresponding point x^i. The set of values in B_{n+1} which Z^α takes on is therefore open. This set, because of the form of Z, is the interior of a cone or cones with central lines as elements. Consider a line on the boundary of this cone: to a sequence of central lines approaching the boundary line from within there corresponds at least one sequence of points in P_n, which has a limit point since P_n is compact. But because of the continuity of Z the limit point must correspond to the boundary line in B_{n+1}. Thus, the boundary also belongs to the set. Hence the set of points corresponding to values of Z^α is also closed, and so it is the whole space B_{n+1}.

As the points of B_{n+1} were represented by $Z^\alpha = \pm e^{z^0} f^\alpha(x^i)$, so the central lines can be represented by f^α, $\pm e^{z^0}$ being merely a parameter running along the line. There remains to prove that no two points in P_n correspond to the same central line. Suppose there were two such points, x_1^i and x_2^i, corresponding to f_1^α and f_2^α with f_1^α and f_2^α proportional. The two points can be connected by a continuous curve $x^i(t)$ $(0 \leqq t \leqq 1)$ and it is supposed that no point on $x^i(t)$ separating x_1 and x_2 corresponds also to kf_1^α. This situation can always be obtained, for on any arc there can be but a finite number of such points. Having

an arc which contains at least two, we can select a pair which are not separated by any other such point.

To the curve $x^i(t)$ there corresponds in B_{n+1} a closed cone given by

$$Z^\alpha(x^0, t) = e^{x^0} f^\alpha(x^i(t)) \cdot k(t) = k e^{x^0} f^\alpha(t), \qquad k(t) = 1 - t\left(1 - \frac{f_1^\alpha}{f_2^\alpha}\right),$$

with $f^\alpha(0) = f_1^\alpha$, $f^\alpha(1) = f_2^\alpha$.

Then consider the family of cones

$$Z^\alpha(x^0, t; \lambda) = \lambda e^{x^0}(k f^\alpha(t) - f_1^\alpha) + e^{x^0} f_1^\alpha \qquad (0 < \lambda \leqq 1).$$

This is a family of cones lying inside the original cone, $Z^\alpha(x^0, t; 1)$, and closing down on the generator f_1^α as $\lambda \to 0$. For any particular value, $\bar{\lambda}$, the corresponding cone can be enclosed in a system of line neighborhoods such that the central lines of the line neighborhoods are in one-to-one correspondence with the points of a system of neighborhoods in P_n , and in each the x's are expressible as analytic functions of the Z's. Each of the cones, therefore, has an image curve $x^i(t, \lambda)$ in P_n for which $x^i(0, \lambda) = x_1^i$ but the other end point, $x^i(1, \lambda)$, may vary with λ. In the family, $x^i(t, 1) = x^i(t)$, the originally selected continuous curve.

However, if $\bar{\lambda}$ is a particular value of λ as in the preceding paragraph, there is a range $\bar{\lambda} - \epsilon < \lambda < \bar{\lambda} + \epsilon$ for which the corresponding cones all lie in the system of line neighborhoods enclosing $Z^\alpha(x^0, t; \bar{\lambda})$. For this range of λ, all the image curves have the same second end point [i.e., $x^i(1, \lambda) = x^i(1, \bar{\lambda})$ for $\bar{\lambda} - \epsilon < \lambda < \bar{\lambda} + \epsilon$] since in the neighborhood of that end point there is but one point corresponding to the generator f_1^α and $Z^\alpha(x^0, 1; \lambda) = e^{x^0} f_1^\alpha$ for all λ.

For the original cone, $Z^\alpha(x^0, t; 1)$, the second end point of the image curve $x^i(t)$ is x_2^i . However, about the generator f_1^α there exists a line neighborhood N whose central lines are in one-to-one correspondence with the points of a neighborhood N_1 about x_1^i , not including x_2^i. A value ϵ can be found such that for $\lambda < \epsilon$ all the cones $Z^\alpha(x^0, t; \lambda)$ lie in the line neighborhood N, and hence their image curves lie entirely in the neighborhood N_1 . It follows from the unique correspondence that the second end point of this set of image curves is x_1^i , and so these curves do not join x_1^i and x_2^i .

There exists, then, a first value $\bar{\lambda}$ of λ, such that $x^i(t, \lambda)$ joins x_1^i and x_2^i for all $\lambda > \bar{\lambda}$; but for any $\eta > 0$, and $\lambda > \bar{\lambda} - \eta$ some image curves do not join x_1^i and x_2^i . Hence, for $\bar{\lambda} - \eta < \lambda < \bar{\lambda} + \eta$, $x^i(1, \lambda)$ is not independent of λ. But this is contradictory to the result just established that for any value of λ there exists a range about it such that the end points are independent of λ in that range.

Thus, no two points x correspond to the same central line in B_{n+1} , and the set of solutions is shown to have the property 2a. The solutions described, therefore, form a set of preferred coördinate systems for projective geometry, extending to the entire space.

PRINCETON UNIVERSITY.

THE WARING PROBLEM WITH SUMMANDS x^m, $m \geqq n$

By Mary Haberzetle

In the original Waring problem the summands considered are positive integral n-th powers. By using different types of summands one obtains various modifications of the problem. In this paper those of the form x^m, $m \geqq n$, are used. If a positive integer is a sum of u such summands with $x \geqq 0$, it is said to be a sum of u "values". If $n \geqq 9$ and $r \leqq 2^n - k - 3$, $I = 2^n + k - 1$ values are necessary and sufficient to represent every positive integer; r is defined by $3^n = 2^n q + r$, $q = [(\frac{3}{2})^n]$, and $k = [.584963n]$. The notation $[x]$ denotes the largest integer $\leqq x$. For $9 \leqq n \leqq 400$ it is shown that the inequality $r \leqq 2^n - k - 3$ is satisfied.

This problem was suggested by L. E. Dickson.

1. **Representation of integers $\leqq 2^n q$.** Write $3^n = 2^n q + r$, $1 \leqq r < 2^n$, $q = [(\frac{3}{2})^n]$.

Lemma A. *Every positive integer $< 2^x - 1$ is a sum of at most $x - 1$ terms* $1, 2, \cdots, 2^{x-1}$.

The lemma may be verified for $x = 1$, $x = 2$. Assume, therefore, that it is true for integers $< 2^x - 1$ and prove by induction. The integer $2^x - 1$ is equal to $1 + 2 + \cdots + 2^{x-1}$ and hence requires x terms. The integers $2^x + 1$, $2^x + 2, \cdots, 2^{x+1} - 2$ are each the sum of 2^x and an integer $< 2^x - 1$. Hence all $< 2^{x+1} - 1$ are sums of at most x terms $1, 2, \cdots, 2^x$.

Let k be the maximum positive integer such that $2^{n+k} \leqq 2^n q$. Then k is the largest integer for which $2^k \leqq q = [(\frac{3}{2})^n]$, that is,

$$k = [(n \log \tfrac{3}{2})/\log 2] = [.584963n].$$

Lemma 1. *All positive integers $\leqq 2^n q$ are sums of $I = 2^n + k - 1$ values. The integer $2^{n+k} - 1$ requires exactly I values.*

Let $B = 2^n x + y$, $0 \leqq y \leqq 2^n - 1$, $0 \leqq x < q$. Then $x < 2^{k+1} - 1$ and by Lemma A is a sum of at most k terms $1, 2^2, \cdots, 2^k$. Hence B is a sum of $2^n + k - 1$ values. Since $q \leqq 2^{k+1} - 1$, $2^n q$ is a sum of $k + 1$ values. The integer $2^{n+k} - 1$ is equal to $2^n(1 + 2 + \cdots + 2^{k-1}) + 2^n - 1$ and hence is a sum of $2^n + k - 1$ values and not fewer.

2. **Results from the asymptotic theory.** The following theorem, whose proof is due to Vinogradow[1] and L. E. Dickson,[2] is stated here with an improved

Received June 9, 1938.

[1] Vinogradow, *On Waring's problem*, Annals of Mathematics, vol. 36(1935), pp. 395–405.

[2] L. E. Dickson, *Proof of the ideal Waring theorem for exponents* 7–180, American Journal of Mathematics, vol. 58(1936), p. 525.

value for N. The improvement is due to some later refinements[3] made in the theory by Dickson.

THEOREM 1. *Let k_0 be the least integer exceeding $-\log r_1/[\log(1-\nu)]$, where*

$$r_1 = \frac{n^2(6n-1)}{n-d}, \qquad d = 1 + 2n^2z, \qquad z = \tfrac{1}{12}\nu^2, \qquad \text{and} \qquad \nu = n^{-1}.$$

Take $k_1 = 2k_0$, $\log_{10} N = .8n^6$. If $n \geqq 9$, every integer $\geqq N$ is a sum of $4n - 2 + 3k_1$ integral n-th powers $\geqq 0$.

3. **Ascent and solution for $9 \leqq n \leqq 400$.** Since $4n - 2 + 3k_1 < 2^n + k$, it follows from Theorem 1 and Lemma 1 that all positive integers $\leqq 2^n q$ and $\geqq N$ are sums of $I = 2^n + k - 1$ values. The following theorem will be proved by ascent.

THEOREM 2. *For $9 \leqq n \leqq 400$ all positive integers are sums of $I = 2^n + k - 1$ values.*

The following lemmas are required.

LEMMA 2. *All integers in the interval $(2^n q, 2^n q + 2^n)$ are sums of E values, where $E = \max(k + r, 2^n - r)$.*

Since $2^n q$ is a sum of $k + 1$ values, $2^n q, 2^n q + 1, \cdots, 2^n q + r - 1$ are sums of $k + r$ values. The integers $2^n q + r = 3^n, \cdots, 2^n q + 2^n - 1 = 3^n - r + 2^n - 1$ are sums of $2^n - r$ values. Also $q + 1 \leqq 2^{k+1}$. Hence $2^n q + 2^n$ is a sum of $k + 1$ values.

The interval in Lemma 2 can be extended to $(2^n q, 2^n q + 2^n q)$ by adding on multiples of 2^n, that is, $2^n x$, $1 \leqq x < q$. By application of Lemma A it is seen that at most k additional values are required to make this extension. Hence all in $(2^n q, 2^n q + 2^n q)$ are sums of $E + k$ values. As in the preceding paragraph the integers in $(2^n q + 2^n q, 2^n q + 2^n q + r - 1)$ can be shown to be sums of $k + r$ and hence of E values. The integer $2^n q + 2^n q + r = 2^n q + 3^n$ is a sum of $k + 2 < E + k$ values. This proves

LEMMA 3. *All integers in $(2^n q, 2^n q + 3^n)$ are sums of $E + k$ values.*

LEMMA[4] 4. *Let L, p, n, z, w be positive integers. If all integers in $(L, L + p^n z)$ are sums of m values, then all in $(L, L + p^n(z + w))$ are sums of $m + w$ values.*

This with $z = 1$ and Lemma 3 give

LEMMA 5. *All integers in $(2^n q, 2^n q + 3^n(1 + w))$ are sums of $E + k + w$ values.*

Take $w = [(\frac{4}{3})^n]$. Then $1 + w > (\frac{4}{3})^n$. This proves

LEMMA 6. *All integers in $(2^n q, 2^n q + 4^n)$ are sums of $E + k + [(\frac{4}{3})^n]$ values.*

By induction one obtains

[3] L. E. Dickson, *Universal forms $\sum a_i x_i^n$ and Waring's problem*, Acta Arithmetica, vol. 2 (1937), pp. 186–191.

[4] Dickson, *Proof of the ideal Waring theorem for exponents* 7–180, op. cit., p. 526.

LEMMA 7. *All integers in the interval* $(2^n q, 2^n q + (n+1)^n)$ *are sums of* $E + k + [(\frac{4}{3})^n] + [(\frac{5}{4})^n] + \cdots + [((n+1)/n)^n]$ *values, and hence are sums of* $E + k + (n-2)[(\frac{4}{3})^n]$ *values.*

LEMMA[5] 8. *Let* $v = (1 - l/L_0)/n, v^n L_0 \geqq 1$. *If all integers between* l *and* L_0 *inclusive are sums of* Q *values, then all between* l *and* L_t *inclusive are sums of* $Q + t$ *values, where*

$$\log_{10} L_t = \left(\frac{n}{n-1}\right)^t (\log_{10} L_0 + n \log_{10} v) - n \log_{10} v. \tag{1}$$

For N in Theorem 1 it is desired that $L_t > N$. Hence t is to be determined such that

$$\left(\frac{n}{n-1}\right)^t (\log_{10} L_0 + n \log_{10} v) - n \log_{10} v > .8n^6. \tag{2}$$

Take $l = 3^n$ and $L_0 = (n+1)^n$. Then the interval (l, L_0) is included in the interval of Lemma 7. Also, the inequality $v^n L_0 \geqq 1$ reduces to $v(n+1) \geqq 1$ and hence to $(\frac{1}{2}(n+1))^{n-1} \geqq 3$, an inequality which holds if $n \geqq 4$.

Case $n = 9$. Here $\log_{10} L_0 = 9$, $\log_{10} v = \bar{1}.0457488$. Increase $.8n^6$ to n^6 in (2). Then $t = 120$. Also $k = 5, r = 227$. Hence $E = \max(k + r, 2^n - r) = 285$. By Lemma 7 and the preceding all integers between $2^n q$ and N inclusive are sums of

$$t + E + k + (n-2)[(\tfrac{4}{3})^n] \tag{3}$$

values. Since $[(\frac{4}{3})^9] = 13$, (3) is equal to $501 < I = 516$.

Case $n = 10$. Now $\log_{10} L_0 = 10.413927$, $\log_{10} v = \bar{2}.9999990$ and from (2), $t = 140$. From $k = 5, r = 681$, it follows that $E = 686$. Also $[(\frac{4}{3})^{10}] = 17$. Hence (3) is equal to $967 < I = 1028$.

When $n \geqq 11$, $v = n^{-1}$ to seven decimal places. From the series for $\log(1 + x)$, $\log_{10} L_0 - n \log_{10} n$ is the product of n by $\log_{10}(1 + n^{-1}) > M(n^{-1} - \frac{1}{2}n^{-2})$, $M = .4343$. Then $L_t > N$ if

$$\left(\frac{n}{n-1}\right)^t M(1 - \tfrac{1}{2}n^{-1}) > n^6, \qquad n \geqq 11. \tag{4}$$

If the denominator $2n$ is replaced by 22, (4) will follow from

$$t \log_{10} \frac{n}{n-1} > .3824138 + 6 \log_{10} n.$$

The coefficient of t is the product of M by

$$\log_e \left(1 + \frac{1}{n-1}\right) > \frac{1}{n-1} - \frac{1}{2(n-1)^2} > \frac{1}{n}, \qquad n > 2.$$

[5] L. E. Dickson, *Recent progress on Waring's theorem and its generalization*, Bulletin of the American Mathematical Society, vol. 39(1933), p. 711.

Hence $L_t > N$ if

$$t > M^{-1}n(.3824138 + 6\log_{10} n), \qquad n \geqq 11. \tag{5}$$

In (3) first let $E = 2^n - r$. Then (3) is $\leqq I = 2^n + k - 1$ if

$$\frac{r}{2^n} \geqq F(n), \qquad F(n) = \frac{t+1}{2^n} + \frac{n-2}{q}. \tag{6}$$

Next let $E = k + r$. The like conclusion holds if

$$\frac{r}{2^n} \leqq 1 - F(n) - \frac{k}{2^n}. \tag{7}$$

Cases $n = 11, 12, 13$. Let $n = 11$. From (5), $t = 168$. Then $F(n) = .187170$ and $.187170 \leqq r/2^n \leqq .809900$. Since $F(n)$ decreases when n increases, Theorem 2 will hold for any $n \geqq 11$ for which $r/2^n$ lies between these decimals. Note that $r/2^n$ is the decimal part of $(\frac{3}{2})^n$. The condition holds for $n = 11, 12, 13$ but not for $n = 14$.

Cases $14 \leqq n \leqq 28$. When $n = 14$ the decimal in (5) reduces to .3780043, and it suffices to take $t = 234$. $F(n) = .055580$. Hence $.055580 \leqq r/2^n \leqq .943932$. This condition is satisfied for $n = 14, 15, \cdots, 28$.

Cases $29 \leqq n \leqq 162$. Let $n = 20$. The decimal part in (5) is reduced to .3732056. Take $t = 377$. Then $.005773 \leqq r/2^n \leqq .994216$. This condition is satisfied for $n = 20, \cdots, 162$.

Cases $163 \leqq n \leqq 400$. Let $n = 100$. The decimal in (5) is reduced to .3643871. Take $t = 2847$. $F(n)$ begins with at least 14 zeros, $k/2^n$ begins with 27 zeros, and hence (7) with at least 14 nines. The decimal $r/2^n$ lies between these limits at least as far as $n = 400$.

4. **Solution for $n \geqq 35$.** The results of this section together with the preceding will show that for $n \geqq 9$ every positive integer is a sum of I values except when $r \geqq 2^n - k - 2$. It can be seen from the preceding section that the inequality $r \geqq 2^n - k - 2$ does not hold for $9 \leqq n \leqq 400$. Whether the inequality is ever satisfied has not yet been determined.

Let $q = [(\frac{3}{2})^n], f = [(\frac{4}{3})^n], g = [(\frac{5}{4})^n]$, and $s = f + 2g$.

LEMMA 9. *If all integers in the interval* $(L, L + 2^n)$ *are sums of* m *values, then all in* $(L, L + (n + (n+1)^n))$ *are sums of*

$$m + k + 1 + f + g + [(\tfrac{4}{3})^n] + \cdots + \left[\left(\frac{n+1}{n}\right)^n\right]$$

values.

The interval $(L, L + 2^n)$ can be extended to $(L, L + 3^n)$ by adding on multiples of 2^n, that is, $2^n x$, $1 \leqq x \leqq q$. The final integer so obtained is $L + 2^n(q+1) > L + 3^n$. The integer x is a sum of at most $k + 1$ terms $1, 2, \cdots, 2^k$. Hence in ascending to $L + 3^n$ at most $k + 1$ additional values are required. To complete the proof of Lemma 9 apply Lemma 4 as in the proof of Lemma 7.

LEMMA 10. *If $n \geqq 35$, $L \leqq 3 \cdot 4^n$, and all integers in the interval $(L, L + 2^n)$ are sums of m values, then every integer $\geqq L$ is a sum of $m + s + k - 1$ values.*

Since $4n - 2 + 3k_1$ is less than $s + k$ when $n \geqq 35$ and hence less than $m + s + k$, it follows from Theorem 1 that all integers $\geqq N$ are sums of $m + s + k - 1$ values. It will be proved that all integers $\geqq L$ and $\leqq N$ are sums of $m + s + k - 1$ values by ascent. By applying Lemma 8 for $n \geqq 35$ as in §3 for $n \geqq 11$, but now with $l = 3 \cdot 4^n$, $L_0 = (n + 1)^n$, and Lemma 9 instead of Lemma 7, it suffices to show that

$$t + 2 + [(\tfrac{9}{8})^n] + \cdots + \left[\left(\frac{n+1}{n}\right)^n\right] \leqq [(\tfrac{5}{4})^n] \tag{8}$$

where $t > M^{-1}n(.3684214 + 6 \log_{10} n)$.

Multiply (8) by $(\frac{4}{5})^n$. The resulting inequality need be verified only for $n = 35$ since np^n, $np^n \log_{10} n$, and p^n decrease when n increases and $n \geqq 7$, $0 < p \leqq \frac{4}{5}$. Take $t = 777$. Then

$$(\tfrac{4}{5})^n\left(t + 2 + [(\tfrac{9}{8})^n] + \cdots + \left[\left(\frac{n+1}{n}\right)^n\right]\right) < .9224296 < [(\tfrac{5}{4})^n](\tfrac{4}{5})^n, \quad n \geqq 35$$

Hence (8) is satisfied.

LEMMA 11. *If $n \geqq 35$ and $s \leqq r \leqq 2^n - k - s$, every positive integer is a sum of I values.*

The integers $2^n q, \cdots, 2^n q + r - 1$ are sums of $k + r \leqq 2^n - s$ values. The integers $\geqq 2^n q + r = 3^n$ and $< 2^n(q + 1)$ and hence $\leqq 3^n + 2^n - r - 1$ are sums of $2^n - r \leqq 2^n - s$ values. Finally, $2^n(q + 1)$ is a sum of $k + 1$ values. Apply Lemma 10 with $L = 2^n q$, $m = 2^n - s$ to obtain the result that every integer $\geqq 2^n q$ is a sum of I values. Apply Lemma 1.

The following lemma is required in the proof of Lemma 13.

LEMMA 12. *If $1 \leqq w \leqq (2^n - 1)r^{-1}$, all integers in $(3^n w, 3^n w + 3^n)$ are sums of M_1 values, where M_1 is the maximum of $A = 2^n - wr + k + w - 1$ and $B = w(k + 1 + r) + k$.*

Those integers $\leqq w3^n + 2^n - wr - 1$ are sums of $w + 2^n - wr - 1$ values. The next integer is equal to $2^n(qw + 1)$. To this add $1, 2, \cdots, wr - 1$ and obtain $3^n w + 2^n - 1$ as the final integer. Therefore all those $\geqq 2^n(qw + 1)$ and $\leqq 3^n w + 2^n$ are sums of $w(k + 1 + r)$ values.

If we add $2^n x$, $1 \leqq x < q$, to each integer in $(3^n w, 3^n w + 2^n)$, the interval is extended to $3^n w + 2^n q$. At most k additional values are required in making the extension. Hence all in $(3^n w, 3^n w + 2^n q)$ are sums of M_1 values. The integers in $(3^n w + 2^n q, 3^n w + 3^n)$ require at most $w + k + r < B$ values.

LEMMA 13. *If $n \geqq 35$ and $r < s$, every positive integer is a sum of I values.*

Since 3 belongs to the exponent 2^{n-2} modulo 2^n, $r = 1$ would imply that n is divisible by 2^{n-2}, whence $n \leqq 4$. If $r = 3$, then $3^{n-1} = (\frac{1}{2}2^n q) + 1$ and $n - 1 \leqq 2$. Hence $r \geqq 5$.

For $x = 0, 1, \cdots, u(3^n - 2^n q) - 1 = ur - 1$, $2^n qu + x$ is a sum of $(k+1)u + x \leqq (k+1)u + ur - 1$ values. If $2^n > ur$, there exist integers $\geqq 3^n u$ and $\leqq 2^n qu + 2^n - 1 = 3^n u + 2^n - 1 - ur$ and each is a sum of $u + 2^n - 1 - ur$ values. The latter and $(k+1)u + ur - 1$ will be $\leqq 2^n - s$ if

$$\frac{s}{r-1} \leqq u \leqq \frac{2^n - s}{k+1+r}. \tag{9}$$

Decrease r to 5 on the left and increase r to s on the right to obtain the subinterval,

$$\frac{s}{4} \leqq u \leqq \frac{2^n - s}{k+1+s}. \tag{10}$$

Thus any u satisfying this inequality will satisfy (9). Since $k + 1 + s > s > r$,

$$\frac{2^n - s}{k+1+s} < \frac{2^n}{r}.$$

Hence, $2^n > ur$.

There will exist an integer u satisfying (10) if the difference between the limits is $\geqq 1$, and therefore, if $2^n \geqq R$,

$$R = 2s + \tfrac{1}{4}s^2 + (1 + \tfrac{1}{4}s)(k+1).$$

But $k + 1 \leqq .584963n + 1$ and $s < 3(\frac{4}{3})^n$. Hence $2^n \geqq R$ if

$$4 \cdot 2^n \geqq 24(\tfrac{4}{3})^n + 9(\tfrac{18}{9})^n + 4(k+1) + 3(\tfrac{4}{3})^n(k+1).$$

Multiply both sides by $(\frac{9}{16})^n$. Then

$$4(\tfrac{9}{8})^n \geqq 9 + 24(\tfrac{3}{4})^n + 4(\tfrac{9}{16})^n(k+1) + 3(\tfrac{3}{4})^n(k+1).$$

Since the right side decreases as n increases, it suffices to verify the inequality for a particular $n \leqq 35$. For $n = 20$ it becomes $42.196 > 9.^+$ and hence holds for $n \geqq 20$.

Thus all integers in $(2^n qu, 2^n qu + 2^n)$ are sums of $2^n - s$ values. By Lemma 10 every integer $\geqq 2^n qu$ is a sum of I values.

It remains to show that every positive integer $< 2^n qu$ is a sum of I values. By Lemma 1 and the proof of Lemma 2 this is true for integers $\leqq 3^n$. It remains then to prove it for integers between 3^n and $3^n U$, where U is the least integer satisfying (10). Lemma 12 is made use of.

For $w \leqq U - 1$, it can be shown that $B \leqq A$. Hence, apply Lemma 12 with M_1 taken equal to A. If $2^n - wr + k + w \leqq 2^n + k$, and hence if $r \geqq 1$, $A \leqq I$. Therefore all integers in $(3^n w, 3^n(w+1))$ are sums of I values. Let $w = 1, 2, \cdots, U - 1$. The proof of Lemma 13 is now complete.

LEMMA 14. *If $n \geqq 35$ and $2^n - k - s < r \leqq 2^n - k - 3$, every positive integer is a sum of I values.*

For $x = 0, 1, \cdots, 3^n u - 2^n(uq + u - 1) - 1 = 2^n - u(2^n - r) - 1$,

$2^n(uq + u - 1) + x$ is a sum of $C = 2^n - u(2^n - r - k - 2) - 2$ values. The integers $\geqq 3^n u$ and $\leqq 2^n(uq + u) - 1 = 3^n u + u(2^n - r) - 1$ are sums of $D = u + u(2^n - r) - 1$ values. Then C and D are each $\leqq 2^n - s$ if

$$\frac{s}{2^n - r - k - 2} \leqq u \leqq \frac{2^n - s}{2^n - r + 1}. \tag{12}$$

Increase r to $2^n - k - 3$ on the left and decrease r to $2^n - k - s + 1$ on the right to obtain the subinterval,

$$s \leqq u \leqq \frac{2^n - s}{k + s}. \tag{13}$$

Note that the greatest x is positive if $u \leqq (2^n - 1)/(2^n - r)$; this follows from (12). There will exist an integer u satisfying (13) if the difference between the limits is $\geqq 1$, and hence if

$$2^n > 2s + ks + s^2 + k.$$

Since $s < 3[(\frac{4}{3})^n]$, the above inequality is true if $2^n > 6(\frac{4}{3})^n + 3(\frac{4}{3})^n k + 9(\frac{16}{9})^n + k$. Multiply each side by $(\frac{9}{16})^n$ to get $(\frac{9}{8})^n > 6(\frac{3}{4})^n + 3(\frac{3}{4})^n k + 9 + k(\frac{9}{16})^n$. This holds for $n \geqq 20$. Thus all integers in $(2^n(uq + u - 1), 2^n(uq + u - 1) + 2^n)$ are sums of $2^n - s$ values. It now follows from Lemma 10 that all integers $\geqq 2^n(uq + u - 1)$ are sums of I values.

It will next be shown that the latter is true also of the integers in

$$(3^n w, 3^n(w + 1)) \text{ for } w = 1, 2, \cdots, U - 1,$$

where U is the least integer satisfying (13) and hence is equal to s. Then it will have been proved that all integers $\geqq 3^n$ are sums of I values. Since all $< 3^n$ are by Lemma 1 and the proof of Lemma 2 sums of I values, the proof of Lemma 14 will be complete.

The inequality

$$2^n \geqq wk + ws + s \tag{14}$$

will be assumed true for the present and later shown to be satisfied for $n \geqq 35$. From (14) it follows that $2^n > w(k + s)$, whence

$$wr > w(2^n - k - s) > (w - 1)2^n, \qquad w(2^n - r) < 2^n.$$

Consider now $J = (3^n w, 3^n w + 2^n)$ and the equations:

	Difference	Weight
$j\{-3^n + (q + 1)2^n\} = j(2^n - r)$	$2^n - r$	jk $(j = 0, 1, \cdots, w - 1)$
$w\{-3^n + (q + 1)2^n\} = w(2^n - r)$	$2^n - w(2^n - r)$	wk.

To difference + weight $- 1$ add w since $w \cdot 3^n$ must be added to get decompositions. For $j \leqq w - 1$ the largest sum is $F = 2^n - r + (w - 1)(k + 1)$. For

the final equation the sum is $G = 2^n - w(2^n - r) + w(k+1) - 1$. Then $G \geqq F$ if $(w+1)r \geqq 2^n w - k$. This follows from (14) and a hypothesis in Lemma 14. Hence every integer in the interval J is a sum of G values. Extend J to be $K = (3^n w, 3^n(w+1))$ by adding $2^n x$, $1 \leqq x \leqq q$, to each integer in J. Then all in K are sums of $G + k + 1$ values. If

$$w(2^n - r - k - 1) \geqq 1, \tag{15}$$

$G + k + 1 \leqq I$. From a hypothesis in Lemma 14, $2^n - r - k - 1 \geqq 2$, and (15) is satisfied.

It remains to show that (14) is true for $n \geqq 35$. Now $w + 1 \leqq U = s < 3(\frac{4}{3})^n$. Hence (14) will follow from $2^n > 3k(\frac{4}{3})^n + 9(\frac{1}{6}\frac{0}{0})^n$. Multiply each side by $(\frac{9}{16})^n$. Then $(\frac{9}{8})^n > 3k(\frac{3}{4})^n + 9$. This holds for $n \geqq 20$.

Lemmas 11 to 14 yield

THEOREM 3. *If $n \geqq 35$ and $r \leqq 2^n - k - 3$, every positive integer is a sum of $I = 2^n + k - 1$ values.*

Let $r \geqq 2^n - k - 2$. If $r = 2^n - 1$, $3^{2n} \equiv 1 \pmod{2^n}$. Since 3 belongs to the exponent 2^{n-2} modulo 2^n, $2n$ would be divisible by 2^{n-2}, and $n \leqq 4$. If $r = 2^n - 3$, $3^{n-1} = (2^n(q+1)/3) - 1$, whence $n \leqq 5$. Since r is odd, it then follows that $r \leqq 2^n - 5$. Thus

$$5 \leqq S \leqq k + 2, \qquad S = 2^n - r. \tag{16}$$

Ascent will be made from an interval containing 4^n. Let

$$4^n = 3^n f + 2^n h + j, \qquad 0 \leqq 2^n h + j < 3^n, \qquad 0 \leqq j < 2^n. \tag{17}$$

Thus $f = [(\frac{4}{3})^n]$. To determine h and j in terms of q, r, and f, replace in (17) 3^n by $2^n q + r$. Then $j + fr \equiv 0, j - fS \equiv 0 \pmod{2^n}$. But $fS \leqq f(k+2) < fq < (\frac{4}{3} \cdot \frac{3}{2})^n$. Hence fS and j are numbers x for which $0 \leqq x < 2^n$. Thus $|j - fS| < 2^n$ and $j = fS$. Then

$$4^n = 2^n(fq + h + f), \qquad h = 2^n - fq - f. \tag{18}$$

If $E_1 = \max(A = 2^n - f(q - S) - 1, 2^n - fS)$, every integer in the interval $(3^n f + 2^n h, 3^n f + 2^n(h+1))$ is a sum of E_1 values as shown by

$$3^n f + 2^n h + x \qquad (x = 0, 1, \cdots, fS - 1),$$

$$4^n + y \qquad (y = 0, 1, \cdots, 2^n - fS - 1).$$

By Lemma 9 and the above all integers in $(4^n, (n+1)^n)$ are sums of $E_1 + k + 1 + [(\frac{4}{3})^n] + \cdots + [((n+1)/n)^n]$ values. Let $n \geqq 35$ and t be the least integer such that $t > M^{-1}n(.3684214 + 6 \log_{10} n)$.

Take $E_1 = 2^n - fS$. Then all integers $\geqq 4^n$ are sums of

$$t + 2^n - fS + k + 1 + f + X \tag{19}$$

values,

$$X = [(\tfrac{5}{4})^n] + \cdots + \left[\left(\frac{n+1}{n}\right)^n\right] < (n-3)(\tfrac{5}{4})^n.$$

If $S - 1 \geqq (t + 2)(\frac{3}{4})^n + (n - 3)(\frac{1}{16})^n$, (19) is $\leqq I$. The right side decreases as n increases if $n > 15$. For $n \geqq 35$, this inequality gives $S \geqq 4.37704$. Since $S \geqq 5$, the number of values is $\leqq I$.

Next let $E_1 = A$. Then all integers $\geqq 4^n$ are sums of $t + A + k + 1 + f + X$ values. This will be $\leqq I$ if $f(q - S - 1) \geqq t + X + 1$, and hence for $n \geqq 35$ if $S \leqq q - 6$. But $S \leqq k + 2 < q - 6$. This proves

THEOREM 4. *If $n \geqq 35$ and $r \geqq 2^n - k - 2$, every integer $\geqq 4^n$ is a sum of $I = 2^n + k - 1$ values.*

UNIVERSITY OF CHICAGO.

ISOMETRIC EMBEDDING OF FLAT MANIFOLDS IN EUCLIDIAN SPACE

By C. Tompkins

1. If $x_1, \cdots, x_{2n}$ represent a set of normal orthogonal coördinates in a Euclidian space of $2n$ dimensions, and if u_a $(a = 1, 2, \cdots, n)$ is a set of parameters which may be reduced modulo 2π, the equations

$$x_{2a-1} = \cos u_a, \qquad x_{2a} = \sin u_a \tag{1}$$

represent an n-dimensional torus (the product of n circles) embedded in the Euclidian space. Furthermore, the metric induced on this torus is flat; with the given choice of parameters it is clear that $g_{ab} = \delta_{ab}$, where g_{ab} represents the fundamental metric tensor induced on the torus and δ_{ab} is zero if a differs from b and one if a and b are equal. This shows that there exists a flat n-dimensional manifold which may be sometrically embedded in $2n$-dimensional Euclidian space.

The object of this note is to show that it is impossible to embed a closed n-dimensional manifold in Euclidian space of less than $2n$ dimensions so that (1) the manifold admits a representation by Frenet equations with coefficients satisfying Gauss equations[1] and (2) the metric induced on the manifold is flat. This result is already known for the case where n is two.[2]

For purposes of convenience, indices chosen from the first part of the alphabet will have the range from 1 to n, and those from the last part have the range 1 to $(n - 1)$. Latin indices, even when repeated, are not summed. Greek indices are summed.

2. **Algebraic lemmas.** Let A be a matrix of n rows and n columns whose elements a_{ab} are vectors in a real space of a fixed number of dimensions. A will be said to satisfy condition I if and only if

$$a_{ac}a_{bd} = a_{ad}a_{bc}, \tag{2}$$

where the multiplication means the scalar product. A will be said to satisfy condition II if and only if the vanishing of any linear combination of any one diagonal element and any set of non-diagonal elements implies that the coefficient of the diagonal element is zero; i.e., each diagonal element is independent of the non-diagonal ones.

Lemma 1. *If the matrix A satisfies both conditions* I *and* II *simultaneously, the elements a_{ab} lie in a space of at least n dimensions.*

Received July 15, 1938. The author is a National Research Fellow.

[1] See H. Weyl, Mathematische Zeitschrift, vol. 12(1922), pp. 154–160; C. Tompkins, Bulletin of the American Mathematical Society, vol. 41(1935), pp. 931–936.

[2] L. P. Eisenhart, *Differential Geometry*, New York, 1909, p. 156.

Proof. Suppose the lemma false. Then there exists a matrix A satisfying conditions I and II, whose elements lie in an $(n-1)$-dimensional space. Since these vectors all lie in an $(n-1)$-dimensional space, at most $(n-1)$ of the vectors a_{1a} are linearly independent. Suppose $\mathfrak{a}$ of them form a maximal set of linearly independent vectors; these $\mathfrak{a}$ can be supposed to be the first $\mathfrak{a}$ vectors of the row, for a_{11} is surely included (because of condition II) among any set of linearly independent vectors of the row and others can be moved to arbitrarily chosen positions on the row by interchanging pairs of columns and the corresponding pairs of rows alternately, a procedure which leaves the matrix satisfying conditions I and II. The vectors in the last $(n-\mathfrak{a})$ positions of the first row can be expressed as linear combinations of those in the first $\mathfrak{a}$ positions; or, what is the same thing, some linear combination of the first $\mathfrak{a}$ columns of the matrix can be added to each of the last $(n-\mathfrak{a})$ so that each of the last $(n-\mathfrak{a})$ elements of the first row becomes zero. Concerning the process of adding one of the first $\mathfrak{a}$ columns to one of the last $(n-\mathfrak{a})$ columns of the matrix, it is important to observe that the process leaves condition I intact and condition II is satisfied by the matrix A_1 with $(n-\mathfrak{a})$ rows and columns in the lower right-hand corner of the new matrix.

Now, condition I states that *every vector of the matrix A_1 is perpendicular to every vector of the first row of A.* For, if a_{ab} is a vector of A_1 (i.e., $\mathfrak{a} < a, b$), condition I states that

$$a_{ab}a_{1c} = a_{ac}a_{1b}. \tag{3}$$

The right member of this equation is zero since $a_{1b} = 0$ if $b > \mathfrak{a}$. Equation (3) states then that a_{ab} is perpendicular to a_{1c} for any allowed choice of indices. This is the assertion of this paragraph.

However, there are supposed to be exactly $\mathfrak{a}$ independent vectors in the first row, so vectors of A_1 must belong to a space of $(n-\mathfrak{a}-1)$ dimensions. A_1 is seen to satisfy both conditions I and II also, and the same reduction can be applied to it and a new and definitely smaller matrix obtained also satisfying the conditions. This leads to an obvious absurdity. Thus the assumption that the proposition is false is not valid.

LEMMA 2. *Let A^z $(z = 1, 2, \cdots, n-1)$ represent a set of $(n-1)$ real symmetric[3] matrices each with n rows and columns and denote by a^z_{ab} the elements of these matrices; suppose that the equations*

$$a^w_{ac}a^w_{bd} = a^w_{ad}a^w_{bc} \tag{4}$$

are satisfied for all choices of indices. Then the determinant of each matrix A^z vanishes.

Proof. If all the matrices A^z are simultaneously multiplied from the left by an orthogonal matrix, and from the right by the transpose of this matrix, a new set of matrices is obtained, and the elements of this set satisfy (4), and the

[3] A. P. Morse has pointed out that these matrices do not have to be symmetric; elementary transformations rather than rotations may be used in the proof.

determinants are the same as the original ones. By such a transformation any of the matrices may be made diagonal.

Now suppose the proposition is false. Then there is a set of matrices A^s satisfying (4), and the determinant of one of these matrices differs from zero. By the procedure outlined above this matrix may be made diagonal, and none of the diagonal elements vanish. Now if we group corresponding elements of all the matrices into vectors, one matrix A can be constructed with elements which are vectors in a space of $(n-1)$ dimensions. Equations (4) state that this matrix satisfies condition I of Lemma 1 and the fact that one of the original matrices was reduced to diagonal form with diagonal elements not zero guarantees that A satisfies condition II. But this, according to Lemma 1, is impossible, for the vectors have only $n-1$ components. Therefore Lemma 2 is true.

3. **The geometric theorem.** In this section the algebraic lemmas of last section will be applied to establish the geometric result of the paper.

THEOREM. *A closed n-dimensional flat manifold which admits a representation by Frenet equations whose coefficients satisfy the Gauss equations can not be embedded isometrically in Euclidian space of less than 2n dimensions.*

Proof. Suppose that the manifold can be embedded in a Euclidian space of $(2n-1)$ dimensions. Then the Frenet equations of the manifold include equations

$$\frac{\partial p_a}{\partial u^b} = \Gamma^a_{ab} p_a + L^\omega_{ab} t_\omega ,$$

where p_a represents the set of n vectors tangent to the manifold which are obtained by differentiating the vector X which expresses the coördinates of the points of the manifold with respect to the n parameters of the manifold, u^a; and t_s represents a set of $(n-1)$ mutually orthogonal unit vectors arbitrarily chosen so that all are perpendicular to the manifold. The coefficients Γ^c_{ab} are the Christoffel symbols of the second kind. If the manifold is flat, the Gauss equations state that

$$L^\omega_{ac} L^\omega_{bd} = L^\omega_{ad} L^\omega_{bc} .$$

Since the manifold is closed, some point on it has a maximum distance from the origin; this point is found by finding the point on the manifold where the function X^2 attains its maximum value, where X again represents the vector from the origin to the point in question and multiplication is in the sense of scalar multiplication of vectors. The theorem will now be proved by demonstrating that such a maximum can not exist under the adopted restriction of a $(2n-1)$-dimensional embedding space.

If X^2 does attain a maximum value, then at that point the scalar product

$$X \frac{\partial X}{\partial u^a} = 0 \tag{5}$$

for every choice of the index a. This implies that at that point the vector X is perpendicular to the manifold; the vectors t_s may then be chosen so that X lies along the first one and perpendicular to all the others. The second derivatives of the coördinate function X exist and are expressed through the Frenet equations; the second derivatives of the function X^2 must also exist then, and the second derivative along any direction of the manifold must certainly be non-positive at the point where X^2 attains its maximum. This means that if u_0^a is the point at which the maximum is attained, and if a^a is any set of n numbers not all zero, and if corresponding to these numbers a curve passing through u_0^a is written

$$u^a = va^a + u_0^a,$$

then

$$\frac{\partial^2 X^2}{\partial v \partial v} \leqq 0$$

at $v = 0$.

Then by a straightforward calculation involving the Frenet equations and equation (5) the second derivative may be computed:

$$\frac{\partial^2 X^2}{\partial v \partial v} = 2g_{\alpha\beta}a^\alpha a^\beta + 2(Xt_\omega)L^\omega_{\alpha\beta}a^\alpha a^\beta.$$

Because of the choice of the vectors t_s, the only one of the scalar products (Xt_s) not zero is (Xt_1), and this is definitely greater than zero. Since the numbers a^a are not all zero,

$$g_{\alpha\beta}a^\alpha a^\beta > 0,$$

for the numbers g_{ab} are coefficients of a positive definite form. Then it follows that

$$L^1_{\alpha\beta}a^\alpha a^\beta < 0, \tag{6}$$

where the choice of a^a is arbitrary except that not all these numbers are zero. Then (6) implies that L^1_{ab} is the matrix of a negative definite form, but the matrices L^s_{ab} satisfy the conditions of Lemma 2 of the preceding section, and therefore the determinant of L^1_{ab} is zero; thus L^1_{ab} can not be the matrix of any definite form and a contradiction is reached.

THE INSTITUTE FOR ADVANCED STUDY.

SIMULTANEOUS INVARIANTS OF A COMPLEX AND SUBCOMPLEX

By C. E. Clark

Part I. Introduction

1. Let P^n (of dimension n) and Q be finite Euclidean polyhedra with $Q \subset P = P^n$. (We use the definitions and notation of P. Alexandroff and H. Hopf, *Topologie*, I.) Moreover, let P and Q admit of a *permissible simplicial division* K_{-1}, i.e., a simplicial division of P such that some subcomplex of K_{-1}, say L_{-1}, is a simplicial division of Q (it is not assumed that the simplexes of K_{-1} are Euclidean simplexes). Let K_1 and L_1 denote the second proper barycentric subdivisions of K_{-1} and L_{-1}, respectively (the word proper indicates that centers of gravity are used). We define the *neighborhood* N_1 of L_1 to be the simplicial complex consisting of the simplexes of K_1 that have at least one vertex in L_1 (together with the sides of all such simplexes).

This neighborhood N_1 depends upon the permissible simplicial division employed. However, we prove in this paper that certain properties of N_1 are independent of the permissible simplicial division.

In the first place, we prove that the structures of the following groups are invariant when the permissible division is changed: (a) the Betti groups (with respect to a pair of coefficient domains as in Alexandroff-Hopf, p. 205) of the simplicial complex B_1 consisting of the simplexes of N_1 that have no vertex in L_1, (b) the Betti groups of the simplicial complex R_1 consisting of the simplexes of K_1 that have no vertex in L_1, (c) groups consisting of homology classes of B_1 that bound in R_1 (cf. $H^r(B_1, R_1)$ of Alexandroff-Hopf, p. 345), and (d) groups consisting of homology classes of B_1 that bound both in R_1 and N_1 (cf. $N^r(N_1, R_1)$ of Alexandroff-Hopf, p. 292, 8). For the definitions of groups (c) and (d) see §8. A relation among the groups is given in §9.

These results have already found applications in the study of generalized and singular manifolds.

The essential feature of the proof is the construction of a set of rays (called deformed rays) which are conveniently related to the subdivision of N_1 and along which cycles can be homotopically deformed.

Part II. The deformed rays

2. **The straight rays.** In §2 we study K_{-1}, a fixed permissible simplicial division of P and Q. We assume in this section that K_{-1} is a Euclidean complex.

Let K_i and L_i ($i = 0, 1, 2, \cdots$) denote the $(i + 1)$-th proper barycentric subdivisions of K_{-1} and L_{-1}, respectively. Let N_i, $i \geqq 0$, be the simplicial

Received August 3, 1938; some of the results of this paper are contained in the thesis presented by the author for the doctor's degree at Cornell University.

complex consisting of the simplexes of K_i that have at least one vertex in L_i (together with the sides of all such simplexes). Let B_i consist of the simplexes of N_i that have no vertex in L_i. We write $B = B_0$, $L = L_0$, $N = N_0$.

We shall first study the regions $N - L$ and construct rays (topological 1-simplexes) such that cycles of $N - L$ can be homotopically deformed by moving the points of the cycles along the rays.

Let σ^p $(p = 1, 2, \cdots, n)$ be a p-simplex of N such that σ^p has at least one vertex in L and at least one vertex not in L. Any such σ^p has a q-dimensional side σ^q in L and the opposite side σ^{p-q-1} in B (because K is the subdivision of K_{-1}).

We know that σ^p is the join (as defined in S. Lefschetz, *Topology*, p. 110) of σ^q and σ^{p-q-1}. Let the straight line segments that join points of σ^q to points of σ^{p-q-1} be called the straight rays of σ^p. The aggregate of all straight rays of all $\sigma^p \subset N$ shall be called the *straight rays of* N. Through any point of $N - (L + B)$ there passes exactly one straight ray of N.

3. $L(\xi)$, $B(\xi)$, **and** $B_1(\xi)$. We shall now use the straight rays of N to project points of $N - (L + B)$ into L, B, and B_1.

For a point $\xi \subset N - (L + B)$ let $U(\xi)$ denote the unique straight ray of N through ξ. Let $L(\xi)$ be the end-point of $U(\xi)$ in L. Let $B(\xi)$ be the end-point of $U(\xi)$ in B. Let $B_1(\xi)$ be the intersection of $U(\xi)$ and B_1.

LEMMA 1. *$L(\xi)$, $B(\xi)$, and $B_1(\xi)$ are single-valued, continuous functions of ξ.*

Proof. The continuity of $L(\xi)$ and $B(\xi)$ is well known.

To prove that $B_1(\xi)$ is single-valued and continuous let $\xi \subset \sigma^p - (\sigma^q + \sigma^{p-q-1})$, σ^p being defined as in §2. We shall represent the points of σ^p in a barycentric coördinate system whose fundamental points are at the vertices of σ^p. Furthermore, coördinates will be normalized so that the sum of the coördinates of a point is always one.

Let τ_j^p $(j = 1, 2, \cdots, p!)$ be the p-simplexes of the proper barycentric subdivision of σ^p that have a vertex at $(1, 0, \cdots, 0)$. Let $\xi = (m_0, m_1, \cdots, m_p)$. We now prove that a necessary and sufficient condition that $\xi \subset \sum_{j=1}^{p!} \tau_j^p$ is $m_0 \geqq m_c$ $(c = 0, 1, \cdots, p)$ (simplexes are closed as point sets). Suppose that $\xi \subset \tau_\nu^p$. We can permute the vertices of the coördinate system so that the coördinates of the vertices of τ_ν^p are $\alpha_0 = (1, 0, \cdots, 0)$, $\alpha_1 = (\frac{1}{2}, \frac{1}{2}, 0, \cdots, 0)$, $\cdots, \alpha_p = \left(\frac{1}{p+1}, \frac{1}{p+1}, \cdots, \frac{1}{p+1}\right)$. In this new coördinate system we have $\xi = (\bar{m}_0, \bar{m}_1, \cdots, \bar{m}_p)$, the $\bar{m}$'s being a permutation of the m's with $\bar{m}_0 = m_0$. Since $\xi \subset \tau_\nu^p$, ξ is the center of gravity of non-negative weights $n_0, n_1, \cdots, n_p, \sum_{c=0}^{p} n_c = 1$, placed respectively at $\alpha_0, \alpha_1, \cdots, \alpha_p$. We have the relations

$$(1) \qquad \bar{m}_c = \frac{1}{c+1} n_c + \frac{1}{c+2} n_{c+1} + \cdots + \frac{1}{p+1} n_p \qquad (c = 0, 1, \cdots, p).$$

Obviously $\bar{m}_0 \geqq \bar{m}_c$ and hence $m_0 \geqq m_c$ $(c = 0, 1, \cdots, p)$.

Conversely, if $m_0 \geqq m_c$, permute the fundamental points of the coördinate system so that $\xi = (\bar{m}_0, \cdots, \bar{m}_p)$, the $\bar{m}$'s being a permutation of the m's with $\bar{m}_0 = m_0$ and $\bar{m}_d \geqq \bar{m}_{d+1}$, $d < p$. In this new coördinate system the points $\alpha_0 = (1, 0, \cdots, 0)$, $\alpha_1 = (\frac{1}{2}, \frac{1}{2}, 0, \cdots, 0), \cdots, \alpha_p = \left(\frac{1}{p+1}, \cdots, \frac{1}{p+1}\right)$ are the vertices of some τ_j^p, say τ_s^p.

Equations (1) can be solved for $n_d = (d+1)(\bar{m}_d - \bar{m}_{d+1}) \geqq 0$, $d < p$, and $n_p = (p+1)\bar{m}_p \geqq 0$. Hence ξ is the center of gravity of non-negative weights placed at the vertices of τ_s^p. Hence $\xi \subset \sum_{j=1}^{p+1} \tau_j^p$.

By this method we prove the

AUXILIARY LEMMA. *The condition $m_a \geqq m_c$ $(c = 0, 1, \cdots, p)$ is a necessary and sufficient condition that $\xi = (m_0, m_1, \cdots, m_p)$ is in a p-simplex τ^p of the proper barycentric subdivision of σ^p such that one vertex of τ^p is the vertex of σ^p associated with the weight m_a.*

We observe that $\xi \subset B_1$ means that ξ is contained in two p-simplexes of the subdivision of σ^p, one of these simplexes having a vertex at a vertex of σ^q and the other having a vertex at a vertex of σ^{p-q-1}. From now on let the first $q+1$ fundamental points of our coördinate system be at the vertices of σ^q. Then $\xi \subset B_1$ if and only if

$$m_\lambda = m_\mu \geqq m_c \qquad (c = 0, 1, \cdots, p) \tag{2}$$

for at least one pair of integers λ, μ such that $0 \leqq \lambda \leqq q < \mu \leqq p$. Let the end-points of the straight ray $U(\xi)$ be $L(\xi) = (l_0, l_1, \cdots, l_q, 0, \cdots, 0)$ and $B(\xi) = (0, \cdots, 0, b_{q+1}, \cdots, b_p)$. The points of $U(\xi)$ are

$$\eta(t) = \{tl_0, tl_1, \cdots, tl_q, (1-t)b_{q+1}, \cdots, (1-t)b_p\} \qquad (0 \leqq t \leqq 1). \tag{3}$$

By (2) the intersections of $U(\xi)$ and B_1 are the points $\eta(t)$ given by the values of t satisfying

$$t \max l_e = (1-t) \max b_f \qquad (0 \leqq e \leqq q < f \leqq p). \tag{4}$$

Since $\max l_e + \max b_f > 0$, $B_1(\xi)$ is single-valued. Since the l's and b's are continuous functions of the coördinates of ξ (using the well-known fact that $L(\xi)$ and $B(\xi)$ are continuous), and since the maximum of a finite set of continuous functions is continuous, it follows from (3) and (4) that $B_1(\xi)$ is continuous in $\sigma^p - (\sigma^q + \sigma^{p-q-1})$. The continuity in $N - (L + B)$ follows immediately.

4. **The point set β.** It happens that the straight rays of N do not always meet B_2 in a single point. However, we shall deform the rays so that they will meet B_2 and determine a single-valued, continuous function.

We now define a point set β such that the straight rays of N meet β in one point and determine a continuous function. Later we shall define a homeomorphism which sends β to B_2.

Let $0 < \epsilon < 1$. Let $E_1 \subset L$ be one end-point of a straight ray U of N and let $E_2 \subset B$ be the other end-point. Let ζ be the unique point of U such that

$$\frac{d(\zeta, E_1)}{d(\zeta, E_2)} = \frac{\epsilon}{1-\epsilon},$$

the symbol d denoting Euclidean distance. Let $\beta(\epsilon)$ be the locus of ζ as U varies in N. Since the complexes considered are finite, we can choose ϵ_1 so small that $\beta(\epsilon_1) \subset N_2$. Let $\beta = \beta(\epsilon_1)$.

Let $\xi = (m_0, m_1, \cdots, m_p) \subset \sigma^p$, where our coördinate system still has its first $q+1$ fundamental points at the vertices of σ^q. ξ is the center of gravity of the two weights $\sum_{d=0}^{q} m_d$ placed at $E_1 = L(\xi)$ and $\sum_{e=q+1}^{p} m_e$ placed at $E_2 = B(\xi)$. Hence we have the

AUXILIARY LEMMA. *A necessary and sufficient condition that $\xi \subset \beta$ is that* $\sum_{e=q+1}^{p} m_e = \epsilon_1$.

For $\xi \subset N - (L + B)$ let $\beta(\xi)$ be the intersection of $U(\xi)$ and β.

LEMMA 2. *$\beta(\xi)$ is single-valued and continuous.*

Proof. $U(\xi)$ is given by (3). By the auxiliary lemma of this section the condition that $\eta(t) \subset \beta$ is $\sum_{f=q+1}^{p} (1-t)b_f = \epsilon_1$. But $\sum_{f=q+1}^{p} b_f = 1$ because the b's are the coördinates of a point. Hence the condition reduces to $t = 1 - \epsilon_1$. Hence $\beta(\xi)$ is unique, and we see from (3) that it is continuous in $\sigma^p - (\sigma^q + \sigma^{p-q-1})$ and hence in $N - (L + B)$.

Consider now the straight rays of N_1 defined for N_1 as the straight rays of N are defined for N. For $\xi \subset N_1 - (L + B_1)$ let $\beta_1(\xi)$ be the intersection of β and the unique straight ray $U_1(\xi)$ of N_1 through ξ.

LEMMA 3. *$\beta_1(\xi)$ is single-valued and continuous in $N_1 - (L + B_1)$.*

Proof. Let $\xi \subset N_1 - (L + B_1)$ and $\xi \subset \sigma^p$ with σ^p still a simplex of $\overline{N - L}$. As before, use a coördinate system with fundamental points at the vertices of σ^p. $U_1(\xi)$ is the join of $(l_0, \cdots, l_q, 0, \cdots, 0)$ and $(n_0, n_1, \cdots, n_p) \subset B_1$. The points of $U_1(\xi)$ are $\zeta(t) = \{tl_0 + (1-t)n_0, \cdots, (1-t)n_{q+1}, \cdots, (1-t)n_p\}$ $(0 \leqq t \leqq 1)$. The condition that $\zeta(t) \subset \beta$ is $\sum_{f=q+1}^{p} (1-t)n_f = \epsilon_1$. We have $\sum_{f=q+1}^{p} n_f > 0$ because otherwise $(n_0, \cdots, n_p)$ would be in L. Hence we can solve for t. To have $0 < t < 1$ we must have $\sum_{f=q+1}^{p} n_f > \epsilon_1$. This inequality is a consequence of the auxiliary lemma of this section and the fact that a straight ray of N meets B, B_1, β, and L in that order. It follows that $\beta_1(\xi)$ is defined and is unique for each ξ; also it is seen that $\beta_1(\xi)$ is continuous.

5. **The homeomorphism $D(\xi)$.** We now define a transformation $D(\xi)$ of $N_1 - (L + B_1)$ into itself. Let U_1 denote a straight ray of N_1 and let P_1 and P_2

be the intersections of the ray U_1 with β and B_2, respectively. Consider the projective transformation of U_1 into itself which leaves its end-points invariant and which carries P_1 into P_2. As U_1 varies in $\overline{N_1 - L}$, we have a (1, 1) transformation of $\overline{N_1 - L}$ into itself which we denote by $\eta = D(\xi)$.

LEMMA 4. *$D(\xi)$ is a homeomorphism of $\overline{N_1 - L}$ into itself.*

Proof. Since $D(\xi)$ is (1, 1), it is sufficient to prove that $D(\xi)$ is continuous in $\overline{N_1 - L}$.

Let $\xi \subset N_1 - (L + B_1)$. Let E_1 and E_2 be the end-points of U_1 in L and B_1, respectively. Then $D(\xi)$ is given by the relation

$$\frac{d(\xi, E_1)}{d(\eta, E_1)} : \frac{d(\xi, E_2)}{d(\eta, E_2)} = \frac{d(P_1, E_1)}{d(P_2, E_1)} : \frac{d(P_1, E_2)}{d(P_2, E_2)}, \tag{5}$$

where P_1, P_2, E_1, and E_2 are determined by $U_1(\xi)$. We rearrange (5) and define f by

$$\frac{d(\eta, E_2)}{d(\eta, E_1)} = \frac{d(\xi, E_2)\, d(P_1, E_1)\, d(P_2, E_2)}{d(\xi, E_1)\, d(P_2, E_1)\, d(P_1, E_2)} = f. \tag{6}$$

We see that $f > 0$ and that f is a continuous function of ξ in $N_1 - (L + B_1)$ (Lemma 1 applied to the straight rays of N_1 and Lemma 3). (6) reduces to

$$d(\eta, E_1) = \frac{d(E_1, E_2)}{1 + f}. \tag{7}$$

We see that $d(\eta, E_1)$ is a continuous function of ξ, and hence η is a continuous function of ξ in $N_1 - (L + B_1)$.

To prove continuity in $\overline{N_1 - L} - B_1$ we observe from (6) and (7) that $d(\eta, E_1)$ approaches zero as ξ approaches E_1. Hence $d(\eta, E_1)$ approaches zero as ξ approaches L because the angles that the straight rays make with any σ^q are greater than a fixed positive number. Since $d(\xi, \eta) < d(\eta, E_1)$ by the definition of $D(\xi)$, it follows that $d(\xi, \eta)$ approaches zero as ξ approaches L. This shows that $D(\xi)$ is continuous in $\overline{N_1 - L}$ at points of L. Finally, we can prove continuity at points of B_1 in a similar way.

6. **The deformed rays.** Apply $D(\xi)$ to the points of the straight rays of N that are in N_1. Let the transforms of the straight rays be called *the deformed rays of N* and be denoted by $\underline{U}(\xi)$.

LEMMA 5. *$\underline{U}(\xi)$ meets L, B, B_1, and B_2 in points $\underline{L}(\xi)$, $\underline{B}(\xi)$, $\underline{B}_1(\xi)$, and $\underline{B}_2(\xi)$ which are single-valued, continuous functions of ξ in $N - (L + B)$.*

This lemma follows from the single-valuedness and continuity of $L(\xi)$, $B(\xi)$, $B_1(\xi)$, and $\beta(\xi)$ and the fact that $\underline{L}(\xi) = LD^{-1}(\xi)$ $(= DLD^{-1}(\xi))$, $\underline{B}(\xi) = BD^{-1}(\xi)$, $\underline{B}_1(\xi) = B_1D^{-1}(\xi)$, and $\underline{B}_2(\xi) = D\beta D^{-1}(\xi)$.

7. **Fundamental theorems.** Let there be given an Abelian group I. All cycles mentioned throughout the paper have coefficients in I (as in Alexandroff-Hopf, p. 176).

Let R_i, $i \geqq 0$, be the simplicial complex consisting of the simplexes of K_i that have no vertex in L.

THEOREM 1. *For $1 \leqq i < j$ there is a homeomorphism between B_i and B_j such that corresponding continuous cycles can be homotopically deformed into each other along deformed rays; furthermore, if the continuous cycles $Z_i \subset B_i$ and $Z_j \subset B_j$ correspond in the homeomorphism, then Z_i bounds in either B_i, N_i, or R_i with respect to the Abelian group $I' \supset I$ if and only if Z_j bounds in the corresponding B_j, N_j, or R_j with respect to I'. To bound with respect to I' is to bound an algebraic complex with coefficients in I'.*

Proof. Just as we have deformed rays of N, we also have deformed rays of N_i, $i > 0$, which meet B_{i+1} and B_{i+2} so as to give single-valued, continuous functions. The deformed rays of N_{i-1}, $i \geqq 1$, determine a homeomorphism between B_i and B_{i+1} such that corresponding continuous cycles can be homotopically deformed into each other along the deformed rays. Let Z_i and Z_{i+1} correspond in this homeomorphism. Since Z_i and Z_{i+1} are homotopic, we have $Z_i \sim Z_{i+1}$ in $\overline{N_i - N_{i+1}}$ with respect to I and hence with respect to $I' \supset I$ (Alexandroff-Hopf, p. 340). Suppose Z_i bounds in X_i with respect to I', where X_i is either B_i, N_i, or R_i. Then Z_{i+1} bounds an algebraic complex Y in $X_i + \overline{N_i - N_{i+1}} = X_{i+1} + \overline{N_i - N_{i+1}}$, the coefficients of Y being in I'. The points of Y that are in $\overline{N_i - N_{i+1}}$ can be continuously deformed into B_{i+1} along the deformed rays. Hence Z_{i+1} bounds in X_{i+1} with respect to I'. The proof of Theorem 1 is readily completed.

THEOREM 2. *Any continuous cycle of $N_i - (L_i + B_i)$, $i \geqq 0$, can be homotopically deformed along deformed rays into a cycle of either L or any B_j, $j \geqq i$; for any given cycle the deformation used is uniquely defined, and the transform of the sum of two cycles is the sum of the transforms.*

Proof. The given cycle can be homotopically deformed along the deformed rays of N_i into either L, B_i, B_{i+1}, or B_{i+2}. The case $j > i + 2$ can now be handled by Theorem 1.

Let k_{-1} be a second permissible simplicial division of P and Q (defined in §1) and let n_i, b_i, and r_i be defined for k_{-1} as N_i, B_i, and R_i are defined for K_{-1}.

THEOREM 3. *If $N_1 \supset n_c \supset N_d \supset n_e$, then any continuous cycle $Z_d \subset N_d$ can be homotopically deformed in $\overline{n_c - n_e}$ into a cycle z_e of n_e; the deformation used is such that z_e is uniquely determined by Z_d and if Z_d^k is deformed into z_d^k $(k = 1, 2)$, then $Z_d^1 + Z_d^2$ is deformed into $z_e^1 + z_e^2$; furthermore z_e bounds in b_e, n_e, or r_e with respect to $I' \supset I$ if and only if Z_d bounds in the corresponding B_d, N_d, or R_d with respect to I'.*

Proof. We consider Z_d to be a continuous cycle of n_c and obtain the uniquely defined cycle z_e of b_e by the procedure of Theorem 2 (using deformed rays of k rather than K). Obviously the transform of the sum of two cycles is the sum of the transforms. Suppose z_e bounds in x_e with respect to I', x_e being b_e, n_e, or r_e. Since Z_d is homotopic and hence homologous to z_e in $\overline{n_c - n_e}$ with respect to I, Z_d bounds in $X_d + \overline{n_c - n_e}$ with respect to I'. As before it follows that Z_d bounds in X_d. The proof is readily completed.

Part III. Invariant groups

8. **Groups related to B_1.** We now define certain groups and prove them to be topologically invariant. Given four Abelian groups I_1, I_2, I_3, and I, the last being a common subgroup of the first three. Consider the r-cycles of B_i, $i \geqq 1$, with coefficients in I which bound algebraic complexes of R_i with coefficients in I_2. These cycles form a group $G^r(B_i, R_i)$ with respect to addition of cycles. Let $H^r(B_i, R_i)$ be the subgroup of G^r $(= G^r(B_i, R_i))$ made up of the cycles of G^r that bound in B_i with respect to I_1. Let $S^r(B_i, R_i) = G^r - H^r$ $(r = 0, 1, \cdots, n)$ (cf. $H^r(B_i, R_i)$ of Alexandroff-Hopf, p. 345).

If $I_2 \supset I_1$, then S^r is the subgroup of the r-th Betti group of B_i (with respect to I, I_1) composed of those homology classes whose cycles bound in R_i with respect to I_2. Even if $I_2 \not\supset I_1$, S^r is the subgroup of the Betti group composed of those homology classes which contain at least one cycle that bounds in R_i with respect to I_2.

Similarly, let $\underline{G}^r(R_i, N_i)$, $i \geqq 1$, be the group of the r-cycles of B_i that bound both in R_i with respect to I_2 and in N_i with respect to I_3. Let $\underline{H}^r(R_i, N_i)$ be the subgroup consisting of the cycles of $\underline{G}^r$ that bound in B_i with respect to I_1. Let $\Gamma^r(R_i, N_i) = \underline{G}^r - \underline{H}^r$ $(r = 0, 1, \cdots, n)$ (cf. the Nahtzyklen of Alexandroff-Hopf, p. 289).

Let $B^r(B_i)$ be the r-th Betti group of B_i with respect to the coefficient domains I, I_1.

Letting small letters still refer to a second permissible division of P and Q, we have

THEOREM 4. *$B^r(B_1) \cong B^r(b_1)$, $S^r(B_1, R_1) \cong S^r(b_1, r_1)$, and $\Gamma^r(R_1, N_1) \cong \Gamma^r(r_1, n_1)$ $(r = 0, 1, \cdots, n)$.*

Proof. We give only the proof of the invariance of S^r. The other two groups are handled in the same way.

We have $S^r(B_1, R_1) \cong S^r(B_d, R_d)$, $d > 1$, as a consequence of Theorem 1.

There is an integer $c \geqq 1$ such that $N_1 \supset n_c$. To see this we take a homeomorph of k that is a Euclidean complex. Then since k is finite, we can subdivide k until the diameters of the simplexes become less than the distance between L and B_1.

Let $N_1 \supset n_c \supset N_d \supset n_e \supset N_f \supset n_g$. We need only show that $S^r(B_d, R_d) \cong S^r(b_e, r_e)$. Consider the cycles of B_d that bound in R_d. By Theorem 3 these cycles deform into a set of cycles of b_e which classify to determine a subgroup of $S^r(b_e, r_e)$, say T, which is isomorphic to $S^r(B_d, R_d)$. It is now sufficient to show that if a cycle $z_e \subset b_e$ bounds in r_e with respect to I_2, then z_e is homologous to one of the cycles that define T. By Theorem 2 z_e can be homotopically deformed into $Z_d \subset B_d$ in $\overline{N_d - N_f}$. We have $Z_d \sim z_e$ in $\overline{N_d - N_f}$ with respect to I. Since z_e bounds in r_e, it follows as above that Z_d bounds in R_d with respect to I_2. Hence there is a cycle $z_e' \subset b_e$ which is homologous to Z_d in $\overline{n_c - n_e}$ with respect to I; z_e' is one of the cycles that define T. We have $z_e \sim z_e'$ in $\overline{n_c - n_g}$ with respect to I and hence in b_e with respect to I_1. The proof is complete.

9. **A relation among the invariant groups.** Let $S^r(L, K)$ be the factor group $A - B$, where A consists of the r-cycles of L that bound in K with respect to I_2, and B is the subgroup consisting of the cycles of A that bound in L with respect to I_1.

THEOREM 5. *If* $I_1 = I_2 = I_3$, *then*

$$\text{rank } S^r(B_1, R_1) = \text{rank } S^r(L, K) + \text{rank } \Gamma^r(R_1, N_1) \qquad (r = 0, 1, \cdots, n)$$

(*the ranks may be finite or infinite*).

Proof. In this proof all cycles have coefficients in I, all homologies and bounding relations are with respect to I_1, and linear independence is always with respect to the integers.

First assume that the ranks of $S^r(L, K)$ and Γ^r are finite and equal to σ and γ, respectively. Then there is a set of r-cycles $x_1, x_2, \cdots, x_\sigma$ of L which are linearly independent with respect to bounding in L and such that if the r-cycle $z \subset L$ bounds in K, then there is a relation $\alpha^0 x + \alpha^c x_c \sim 0$ in L, the α's being integers with $\alpha^0 \neq 0$ and the repeated index indicating summation. Similarly there is a maximum set of linearly independent r-cycles of B_1 that bound in R_1 and N_1, say $g_1, g_2, \cdots, g_\gamma$.

We have $x_c = \dot{C}$, the dot indicating boundary and C being an algebraic complex of K with coefficients in I_1. We may assume that x_c and C are simplicial complexes of K_1. We have $C = C_1 + C_2$, $C_1 \subset N_1$ and $C_2 \subset R_1$ (Alexandroff-Hopf, p. 289). Let $y_c = x_c - \dot{C}_1 = \dot{C}_2$. We have $y_c \subset B_1$, $y_c \sim x_c$ in N_1, and $y_c \sim 0$ in R_1.

We next show that $y_1, y_2, \cdots, y_\sigma, g_1, g_2, \cdots, g_\gamma$ are independent with respect to bounding in B_1. Suppose $\beta^c y_c + \delta^d g_d \sim 0$ in B_1, the β's and δ's being integers. Since each $g_d \sim 0$ in N_1 by definition, we have $\beta^c y_c \sim 0$ in N_1. Hence $\beta^c x_c \sim 0$ in N_1 and hence in L by Theorem 2. Hence $\beta^c = 0$ $(c = 1, 2, \cdots, \sigma)$. Hence $\delta^d g_d \sim 0$ in B_1. Hence $\delta^d = 0$ $(d = 1, 2, \cdots, \gamma)$ and the desired independence is proved.

It follows that rank $S^r(B_1, R_1) \geqq \sigma + \gamma$. We now show that the equality holds. To do this assume that $z \subset B_1$ bounds in R_1. By Theorem 2 z can be deformed into $z' \subset L$ such that $z' \sim z$ in N_1. We see that z' bounds in K. Hence there are α's such that $\alpha^0 z' + \alpha^c x_c \sim 0$ in L, $\alpha^0 \neq 0$. This implies $\alpha^0 z + \alpha^c y_c \sim 0$ in N_1. But this homology holds in R_1 because each term bounds in R_1. Hence by the definition of the g's we have a relation $\beta^0(\alpha^0 z + \alpha^c y_c) + \beta^d g_d \sim 0$ in B_1, $\beta^0 \neq 0$. This proves that rank $S^r(L, K) = \sigma + \gamma$ and that Theorem 5 holds when σ and γ are finite.

Suppose now that rank $S^r(B_1, R_1)$ is infinite. The above discussion shows that one of the other ranks must be infinite and that the theorem holds. If γ is infinite, then rank $S^r(B_1, R_1)$ is infinite because Γ^r is a subgroup of $S^r(B_1, R_1)$. Finally suppose that σ is infinite. By the above procedure of obtaining the y's from the x's it is easily shown that rank $S^r(B_1, R_1)$ is infinite. Proof of Theorem 5 is complete.

10. We now show by example that Theorem 5 can not be generalized to read $S^r(B_1, R_1) \cong S^r(L, K) + \Gamma^r(R_1, N_1)$. Consider a projective plane π and a 2-simplex σ. Determine a homeomorphism between the boundary of σ and a projective line of π. Identify points that correspond by the homeomorphism. Let L be π. B_1 is a circle on σ. Use the integers for all coefficient domains. Then $S^1(B_1, R_1)$ is the additive group of the integers, and $S^1(L, K)$ is of order 2. Hence the above group isomorphism can not hold regardless of what Γ^1 is.

11. THEOREM 6. *The Betti groups of R and r (with respect to I, I') are isomorphic.*

This result follows from the theorems of §7 by the methods of Part III.

Appendix

If each straight ray of N meets B_2 in exactly one point and if the projection of points into B_2 along the straight rays is a continuous transformation, then it is unnecessary to introduce deformed rays. However, we give here an example of a K and L such that some straight rays meet B_2 in more than one point.

Let σ^p still be a p-simplex with one side in L and the opposite side in B.

THEOREM. *Given $\xi \subset \sigma^p$; a necessary and sufficient condition that $\xi \subset N_2$ is that there is a barycentric coördinate system with fundamental points at the vertices of σ^p such that the following conditions hold:*

(a) *if the coördinates of ξ are $(m_1, m_2, \cdots, m_{p+1})$, then $m_j \geqq m_{j+1}$ $(j \leqq p)$,*

(b) *for some integer $\alpha > 0$ the first α fundamental points are in L while the $(\alpha + 1)$-th fundamental point is in B, and*

(c) *for some integer $\beta \leqq \alpha$ we have*

$$\beta(m_\beta - m_{\beta+1}) \geqq i(m_i - m_{i+1}) \qquad (i \leqq p+1) \tag{8}$$

with the convention that $m_{p+2} = 0$.

Proof. First we prove the necessity of the condition. Let τ be a p-simplex of the barycentric subdivision of σ^p with $\xi \subset \tau$. Consider the coördinate system such that the coördinates of the k-th vertex of τ are the elements of the k-th row of

$$\left\| \begin{array}{ccccc} 1 & 0 & 0 & \cdots & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{p+1} & \frac{1}{p+1} & \frac{1}{p+1} & \cdots & \frac{1}{p+1} \end{array} \right\|. \tag{9}$$

We shall show that for this coördinate system the conditions (a), (b), and (c) hold.

From (9) we see that the first k vertices of τ are in L if and only if the first k fundamental points are in L. Also if the $(k+1)$-th fundamental point is in B, then every vertex after the k-th is not in L. Since $\xi \subset \tau \cdot N_2$, the first

vertex of τ is in L. Hence there is a number $\alpha > 0$ such that the first α fundamental points are in L but the $(\alpha + 1)$-th is in B. Condition (b) is verified.

Since $\xi \subset \tau$, ξ is the center of gravity of a set of non-negative weights $n_1, n_2, \cdots, n_{p+1}$, $\sum_{i=1}^{p+1} n_i = 1$, n_i being placed at the i-th vertex of τ. Since $\xi \subset N_2$ and since the first α and only the first α vertices of τ are in L, it follows from the auxiliary lemma of §3 that there is an integer $\beta \leqq \alpha$ such that $n_\beta \geqq n_i$, $i \leqq p + 1$.

As in §3 we have

$$m_i = \frac{1}{i} n_i + \frac{1}{i+1} n_{i+1} + \cdots + \frac{1}{p+1} n_{p+1} \qquad (i \leqq p + 1). \tag{10}$$

From (10) $m_j \geqq m_{j+1}$, $j \leqq p$. Hence condition (a) holds. Solving (10) we get

$$n_i = i(m_i - m_{i+1}) \qquad (i \leqq p + 1), \tag{11}$$

where $m_{d+2} = 0$. From $n_\beta \geqq n_i$ we get condition (c). The necessity of the condition that $\xi \subset N_2$ is proved.

To prove the sufficiency we start with the hypothesis that a coördinate system is given with fundamental points at the vertices of σ^p and such that conditions (a), (b), and (c) hold for a given point ξ. We must show that $\xi \subset N_2$.

The rows of the matrix (9) are the coördinates of the vertices of a simplex τ of the subdivision of σ^p. Condition (b) states that the first α vertices of τ are in L. Solving (10) we obtain a set of weights $n_1, n_2, \cdots, n_{p+1}$ given by (11). From (11) and condition (a) we see that the n's are non-negative. From (11) and condition (c) we have $n_\beta \geqq n_i$, $i \leqq p + 1$. Using the auxiliary lemma of §3, we infer that ξ is in a p-simplex ω of the subdivision of τ such that a vertex of ω is in L. This means $\xi \subset N_2$. Proof of the theorem is complete.

We now exhibit a straight ray that meets B_2 in more than one point. Consider a σ^{92} with ninety-one vertices in L and two in B. Consider the straight ray that joins $(\frac{1}{1\,0}, \frac{1}{1\,0\,0}, \frac{1}{1\,0\,0}, \cdots, \frac{1}{1\,0\,0}, 0, 0)$ and $(0, \cdots, 0, \frac{2}{3}, \frac{1}{3})$. The points of this ray are

$$\eta(t) = \{\tfrac{1}{1\,0}t, \tfrac{1}{1\,0\,0}t, \cdots, \tfrac{1}{1\,0\,0}t, \tfrac{2}{3}(1-t), \tfrac{1}{3}(1-t)\} \quad (0 \leqq t \leqq 1).$$

Using the above theorem, we verify that $\eta(\frac{1}{1\,0\,0\,3})$ is in N_2, but $\eta(\frac{1}{2\,0\,0\,3})$ is not in N_2. Hence the straight ray passes in and out of N_2 and meets B_2 at least three times.

UNIVERSITY OF ROCHESTER.

LINEAL ELEMENT TRANSFORMATIONS OF SPACE FOR WHICH NORMAL CONGRUENCES OF CURVES ARE CONVERTED INTO NORMAL CONGRUENCES

By Edward Kasner

A doubly infinite family of curves in space for which orthogonal surfaces can be constructed is called a *normal congruence*. In this paper all transformations of lineal elements (x, y, z, y', z') of space are determined such that every normal congruence of curves shall be converted into a normal congruence. The infinite group obtained is shown to be isomorphic with the group of contact transformations in space of planar or surface elements (x, y, z, P, Q). (The only transformations in the new group which convert curves into curves are the conformal transformations, which form a ten-parameter group.) Our result may also be stated in this form: the only transformations which carry every pair of partial differential equations in involution into a pair of partial differential equations in involution are the contact transformations. That is, if every set of ∞^3 planar elements which are obtained from a set of ∞^1 surfaces is sent into a set of the same kind, then necessarily every single union is converted into a union.

We thus obtain a new characterization of the contact group in space. We do not *assume* that the individual surfaces in the family of ∞^1 surfaces are converted into surfaces. But from our complicated proof it does result that if *every* integrable field becomes such a field, then the individual unions are actually converted into individual unions; and *therefore* the result is a contact transformation.

A lineal element E is usually defined by the coördinates (x, y, z, y', z'), where (x, y, z) are the Cartesian coördinates of the point of the element E and $(1, y', z')$ are the direction numbers of the direction of the element E. From this it is of course obvious that in the case where ∞^1 lineal elements form a curve (or union) $y' = dy/dx$ and $z' = dz/dx$. Thus y' is the total derivative of y with respect to x and z' is the total derivative of z with respect to x. But in our work it will be more convenient to define an element E by the coördinates (x, y, z, p, q), where (x, y, z) are the Cartesian coördinates of the point of the element E and $(p, q, -1)$ are the direction numbers of the direction of the element E. From this it is seen that in the case where ∞^1 lineal elements form a curve (or union) $p = -dx/dz$ and $q = -dy/dz$. Thus p is minus the total derivative of x with respect to z and q is minus the total derivative of y with respect to z. The relationships between the old and new coördinate systems are obviously $p =$

Received August 22, 1938. The results were presented before the American Mathematical Society (1906) and the International Congress at Zurich (1932). I wish to thank J. De Cicco for his valuable assistance in writing this paper. A somewhat shorter discussion, using infinitesimal transformations, will be published elsewhere.

$-1/z'$ and $q = -y'/z'$. Henceforth, when we speak of the coördinates of a lineal element, we shall always mean the coördinates (x, y, z, p, q).

A planar element E is, of course, defined by the coördinates (x, y, z, P, Q), where (x, y, z) are the Cartesian coördinates of the point of the planar element E and $(P, Q, -1)$ are the direction numbers of the normal direction of the plane of the planar element E. From this it is seen that, if ∞^2 planar elements form a surface (or union), $P = \partial z/\partial x$, $Q = \partial z/\partial y$. Thus P is the partial derivative of z with respect to x and Q is the partial derivative of z with respect to y.

From the preceding it is an immediate consequence that if a lineal element and a planar element which have a common point (x, y, z) are orthogonal, then they must have the same coördinates; that is, $p = P$, and $q = Q$. The advantage of using the coördinates (x, y, z, p, q) for a lineal element E is now seen as follows. When we find the group of lineal element transformations which convert every normal field of lineal elements into a normal field of lineal elements, we are led to the conclusion that the functions $X = X(x, y, z, p, q)$, $Y = Y(x, y, z, p, q)$, $Z = Z(x, y, z, p, q)$, $P = P(x, y, z, p, q)$, $Q = Q(x, y, z, p, q)$, which define any lineal element transformation of the group, must satisfy the partial differential equations obtained by Lie in connection with the group of contact transformations of planar elements. Since (x, y, z, p, q) also denotes the planar element orthogonal to the lineal element, it is then immediately obvious that the group of lineal element transformations which convert every normal field into a normal field is isomorphic with the group of contact transformations of planar elements.

1. **Normal congruences of curves.** A doubly infinite family of curves in space for which orthogonal surfaces can be constructed is called a *normal congruence.*

It is found that *the necessary and sufficient condition that the* ∞^2 *curves whose differential equations are*

$$\frac{dx}{dz} = -p(x, y, z), \qquad \frac{dy}{dz} = -q(x, y, z)$$

be a normal congruence is that the functions $p(x, y, z)$ *and* $q(x, y, z)$ *satisfy the condition*

$$pq_z - qp_z - p_y + q_x = 0.$$

2. **Normal fields of lineal elements.** A set of ∞^3 lineal elements in space is said to be a *normal field* if the set of ∞^2 curves containing them is a normal congruence.

From §1 and the definition of a normal field of lineal elements, it is found that *the necessary and sufficient condition that the* ∞^3 *lineal elements*

$$p = p(x, y, z), \qquad q = q(x, y, z)$$

be a normal field is that the functions $p(x, y, z)$ *and* $q(x, y, z)$ *satisfy the partial differential equation*

$$pq_z - qp_z - p_y + q_x = 0.$$

3. Normal fields of lineal elements into normal fields of lineal elements.

FUNDAMENTAL THEOREM. *The necessary and sufficient conditions that the lineal element transformation*

$$X = X(x, y, z, p, q), \qquad Y = Y(x, y, z, p, q), \qquad Z = Z(x, y, z, p, q),$$
$$P = P(x, y, z, p, q), \qquad Q = Q(x, y, z, p, q)$$

convert every normal field of lineal elements into a normal field of lineal elements are that the functions X, Y, Z, P, Q satisfy the four partial differential equations of first order

$$\text{(L)} \qquad \begin{aligned} Z_x + pZ_z &= P(X_x + pX_z) + Q(Y_x + pY_z), \\ Z_y + qZ_z &= P(X_y + qX_z) + Q(Y_y + qY_z), \\ Z_p &= PX_p + QY_p, \qquad Z_q = PX_q + QY_q . \end{aligned}$$

Thus the infinite group of lineal element transformations which convert every normal field of lineal elements into a normal field of lineal elements is isomorphic with the group of contact transformations of surface elements.

Under our transformation, we must have

$$\text{(1)} \qquad PQ_z - QP_z - P_Y + Q_X = 0,$$

whenever

$$pq_z - qp_z - p_y + q_x = 0.$$

It is found that (1) may be written in the form

$$\text{(2)} \qquad \begin{aligned} &P \begin{vmatrix} X_x + X_p p_x + X_q q_x & X_y + X_p p_y + X_q q_y & X_z + X_p p_z + X_q q_z \\ Y_x + Y_p p_x + Y_q q_x & Y_y + Y_p p_y + Y_q q_y & Y_z + Y_p p_z + Y_q q_z \\ Q_x + Q_p p_x + Q_q q_x & Q_y + Q_p p_y + Q_q q_y & Q_z + Q_p p_z + Q_q q_z \end{vmatrix} \\ &-Q \begin{vmatrix} X_x + X_p p_x + X_q q_x & X_y + X_p p_y + X_q q_y & X_z + X_p p_z + X_q q_z \\ Y_x + Y_p p_x + Y_q q_x & Y_y + Y_p p_y + Y_q q_y & Y_z + Y_p p_z + Y_q q_z \\ P_x + P_p p_x + P_q q_x & P_y + P_p p_y + P_q q_y & P_z + P_p p_z + P_q q_z \end{vmatrix} \\ &- \begin{vmatrix} X_x + X_p p_x + X_q q_x & X_y + X_p p_y + X_q q_y & X_z + X_p p_z + X_q q_z \\ P_x + P_p p_x + P_q q_x & P_y + P_p p_y + P_q q_y & P_z + P_p p_z + P_q q_z \\ Z_x + Z_p p_x + Z_q q_x & Z_y + Z_p p_y + Z_q q_y & Z_z + Z_p p_z + Z_q q_z \end{vmatrix} \\ &+ \begin{vmatrix} Q_x + Q_p p_x + Q_q q_x & Q_y + Q_p p_y + Q_q q_y & Q_z + Q_p p_z + Q_q q_z \\ Y_x + Y_p p_x + Y_q q_x & Y_y + Y_p p_y + Y_q q_y & Y_z + Y_p p_z + Y_q q_z \\ Z_x + Z_p p_x + Z_q q_x & Z_y + Z_p p_y + Z_q q_y & Z_z + Z_p p_z + Z_q q_z \end{vmatrix} = 0, \end{aligned}$$

where $q_x = p_y + qp_z - pq_z$.

The equation (2) must be an identity in p_x, p_y, p_z, q_y, q_z, after we eliminate q_x. Upon setting the coefficients equal to zero, we thus obtain the nine equations

$$
\begin{vmatrix} \bar{p} & \bar{q} & \bar{z} \\ Y_p & Y_q & Y_z \\ Q_p & Q_q & Q_z \end{vmatrix} + \begin{vmatrix} \bar{p} & \bar{q} & \bar{z} \\ X_p & X_q & X_z \\ P_p & P_q & P_z \end{vmatrix} = 0,
$$

$$
\begin{vmatrix} \bar{p} & \bar{y} & \bar{q} \\ Y_p & Y_y & Y_q \\ Q_p & Q_y & Q_q \end{vmatrix} + \begin{vmatrix} \bar{p} & \bar{y} & \bar{q} \\ X_p & X_y & X_q \\ P_p & P_y & P_q \end{vmatrix} = 0,
$$

$$
\begin{vmatrix} \bar{p} & \bar{y} & \bar{z} \\ Y_p & Y_y & Y_z \\ Q_p & Q_y & Q_z \end{vmatrix} + \begin{vmatrix} \bar{p} & \bar{y} & \bar{z} \\ X_p & X_y & X_z \\ P_p & P_y & P_z \end{vmatrix} = 0,
$$

$$
\begin{vmatrix} \bar{x} & \bar{q} & \bar{p} \\ Y_x & Y_q & Y_p \\ Q_x & Q_q & Q_p \end{vmatrix} + \begin{vmatrix} \bar{x} & \bar{q} & \bar{p} \\ X_x & X_q & X_p \\ P_x & P_q & P_p \end{vmatrix} = 0,
$$

$$
\begin{vmatrix} \bar{x} & \bar{q} & \bar{z} \\ Y_x & Y_q & Y_z \\ Q_x & Q_q & Q_z \end{vmatrix} + \begin{vmatrix} \bar{x} & \bar{q} & \bar{z} \\ X_x & X_q & X_z \\ P_x & P_q & P_z \end{vmatrix} = 0,
$$

$$
\begin{vmatrix} \bar{x} & \bar{y} & \bar{z} \\ Y_x & Y_y & Y_z \\ Q_x & Q_y & Q_z \end{vmatrix} + \begin{vmatrix} \bar{x} & \bar{y} & \bar{z} \\ X_x & X_y & X_z \\ P_x & P_y & P_z \end{vmatrix} = 0, \tag{3}
$$

$$
\begin{vmatrix} \bar{x} & \bar{p} & \bar{z} \\ Y_x & Y_p & Y_z \\ Q_x & Q_p & Q_z \end{vmatrix} + \begin{vmatrix} \bar{x} & \bar{p} & \bar{z} \\ X_x & X_p & X_z \\ P_x & P_p & P_z \end{vmatrix} + \begin{vmatrix} \bar{q} & \bar{y} & \bar{z} \\ Y_q & Y_y & Y_z \\ Q_q & Q_y & Q_z \end{vmatrix} + \begin{vmatrix} \bar{q} & \bar{y} & \bar{z} \\ X_q & X_y & X_z \\ P_q & P_y & P_z \end{vmatrix} = 0,
$$

$$
\begin{vmatrix} \bar{x} & \bar{y} & \bar{p} \\ Y_x & Y_y & Y_p \\ Q_x & Q_y & Q_p \end{vmatrix} + \begin{vmatrix} \bar{x} & \bar{y} & \bar{p} \\ X_x & X_y & X_p \\ P_x & P_y & P_p \end{vmatrix} + q \begin{vmatrix} \bar{q} & \bar{y} & \bar{z} \\ Y_q & Y_y & Y_z \\ Q_q & Q_y & Q_z \end{vmatrix} + q \begin{vmatrix} \bar{q} & \bar{y} & \bar{z} \\ X_q & X_y & X_z \\ P_q & P_y & P_z \end{vmatrix} = 0,
$$

$$\begin{vmatrix} \bar{x} & \bar{y} & \bar{q} \\ Y_x & Y_y & Y_q \\ Q_x & Q_y & Q_q \end{vmatrix} + \begin{vmatrix} \bar{x} & \bar{y} & \bar{q} \\ X_x & X_y & X_q \\ P_x & P_y & P_q \end{vmatrix}$$

$$- p \begin{vmatrix} \bar{q} & \bar{y} & \bar{z} \\ Y_q & Y_y & Y_z \\ Q_q & Q_y & Q_z \end{vmatrix} - p \begin{vmatrix} \bar{q} & \bar{y} & \bar{z} \\ X_q & X_y & X_z \\ P_q & P_y & P_z \end{vmatrix} = 0,$$

where

$$\begin{aligned} \bar{x} &= Z_x - PX_x - QY_x, \\ \bar{y} &= Z_y - PX_y - QY_y, \\ \bar{z} &= Z_z - PX_z - QY_z, \\ \bar{p} &= Z_p - PX_p - QY_p, \\ \bar{q} &= Z_q - PX_q - QY_q. \end{aligned} \tag{4}$$

Our problem is to discuss this complicated set (3) *and to show that we necessarily obtain Lie's conditions* (L). The proof is long, covering the next seven pages. It is easy to show that (3) is a consequence of (L). The converse is difficult. We introduce the following abbreviations:

$$A = \begin{vmatrix} Y_x & Y_q \\ Q_x & Q_q \end{vmatrix} + \begin{vmatrix} X_x & X_q \\ P_x & P_q \end{vmatrix}, \qquad B = \begin{vmatrix} Y_x & Y_p \\ Q_x & Q_p \end{vmatrix} + \begin{vmatrix} X_x & X_p \\ P_x & P_p \end{vmatrix},$$

$$C = \begin{vmatrix} Y_p & Y_q \\ Q_p & Q_q \end{vmatrix} + \begin{vmatrix} X_p & X_q \\ P_p & P_q \end{vmatrix}, \qquad D = \begin{vmatrix} Y_y & Y_q \\ Q_y & Q_q \end{vmatrix} + \begin{vmatrix} X_y & X_q \\ P_y & P_q \end{vmatrix},$$

$$E = \begin{vmatrix} Y_y & Y_p \\ Q_y & Q_p \end{vmatrix} + \begin{vmatrix} X_y & X_p \\ P_y & P_p \end{vmatrix}, \qquad F = \begin{vmatrix} Y_y & Y_x \\ Q_y & Q_x \end{vmatrix} + \begin{vmatrix} X_y & X_x \\ P_y & P_x \end{vmatrix}, \tag{5}$$

$$G = \begin{vmatrix} Y_x & Y_p \\ Q_x & Q_p \end{vmatrix} + \begin{vmatrix} X_x & X_p \\ P_x & P_p \end{vmatrix}, \qquad H = \begin{vmatrix} Y_x & Y_q \\ Q_x & Q_q \end{vmatrix} + \begin{vmatrix} X_x & X_q \\ P_x & P_q \end{vmatrix},$$

$$I = \begin{vmatrix} Y_x & Y_z \\ Q_x & Q_z \end{vmatrix} + \begin{vmatrix} X_x & X_z \\ P_x & P_z \end{vmatrix}, \qquad J = \begin{vmatrix} Y_z & Y_y \\ Q_z & Q_y \end{vmatrix} + \begin{vmatrix} X_z & X_y \\ P_z & P_y \end{vmatrix}.$$

From the equations (5) we derive the following consequences:

$$-AE + BD + CF = \begin{vmatrix} Y_y & Y_x & Y_p & Y_q \\ Q_y & Q_x & Q_p & Q_q \\ X_y & X_x & X_p & X_q \\ P_y & P_x & P_p & P_q \end{vmatrix},$$

$$-AG + BH + CI = \begin{vmatrix} Y_x & Y_z & Y_p & Y_q \\ Q_x & Q_z & Q_p & Q_q \\ X_x & X_z & X_p & X_q \\ P_x & P_z & P_p & P_q \end{vmatrix},$$

$$CJ - DG + EH = \begin{vmatrix} Y_x & Y_y & Y_p & Y_q \\ Q_x & Q_y & Q_p & Q_q \\ X_x & X_y & X_p & X_q \\ P_x & P_y & P_p & P_q \end{vmatrix}, \tag{6}$$

$$AJ - DI + FH = \begin{vmatrix} Y_x & Y_y & Y_z & Y_q \\ Q_x & Q_y & Q_z & Q_q \\ X_x & X_y & X_z & X_q \\ P_x & P_y & P_z & P_q \end{vmatrix},$$

$$BJ - EI + FG = \begin{vmatrix} Y_z & Y_y & Y_z & Y_p \\ Q_z & Q_y & Q_z & Q_p \\ X_x & X_y & X_z & X_p \\ P_x & P_y & P_z & P_p \end{vmatrix}.$$

By means of the equations (5), the nine fundamental equations (3) may be written as

$$\begin{aligned} C\bar{z} - A\bar{p} + B\bar{q} &= 0, \\ -C\bar{y} + D\bar{p} - E\bar{q} &= 0, \\ B\bar{y} - E\bar{z} + F\bar{p} &= 0, \\ -C\bar{x} + H\bar{p} - G\bar{q} &= 0, \\ -A\bar{x} + H\bar{z} - I\bar{q} &= 0, \\ F\bar{x} - I\bar{y} + J\bar{z} &= 0, \\ -B\bar{x} + A\bar{y} + (G - D)\bar{z} - I\bar{p} + F\bar{q} &= 0, \\ E\bar{x} + (-G + qA)\bar{y} - qD\bar{z} + J\bar{p} + qF\bar{q} &= 0, \\ D\bar{x} + (-H - pA)\bar{y} + pD\bar{z} + (J - pF)\bar{q} &= 0. \end{aligned} \tag{7}$$

We shall prove first that $C = 0$. For let us suppose that $C \neq 0$. Then the equations (7) are equivalent to the equations

$$\bar{x} = \frac{H}{C}\bar{p} - \frac{G}{C}\bar{q}, \qquad \bar{y} = \frac{D}{C}\bar{p} - \frac{E}{C}\bar{q}, \qquad \bar{z} = \frac{A}{C}\bar{p} - \frac{B}{C}\bar{q},$$

$$\begin{aligned}
&(-AE + BD + CF)\bar{p} = 0,\\
&(-AG + BH + CI)\bar{q} = 0,\\
&(-AG + BH + CI)\bar{p} - (-AE + BD + CF)\bar{q} = 0,\\
&(AJ - DI + FH)\bar{p} - (BJ - EI + FG)\bar{q} = 0,\\
&(CJ - DG + EH)\bar{p} + q(-AE + BD + CF)\bar{q} = 0,\\
&[(CJ - DG + EH) - p(-AE + BD + CF)]\bar{q} = 0.
\end{aligned} \tag{8}$$

It is seen that $\bar{p}$ and $\bar{q}$ cannot both be zero. For then from (8) we find that $\bar{x} = \bar{y} = \bar{z} = \bar{p} = \bar{q} = 0$, and hence by (4) the Jacobian of our transformation is zero. This proves that at least one of the quantities $\bar{p}$ and $\bar{q}$ is not zero.

In the fourth, fifth, sixth, eighth, ninth of the equations (8), we find that the conditions that at least one of the quantities $\bar{p}$ and $\bar{q}$ be different from zero are

$$\begin{aligned}
-AE + BD + CF &= 0,\\
-AG + BH + CI &= 0,\\
CJ - DG + EH &= 0.
\end{aligned} \tag{9}$$

From the fact that $C \neq 0$, and from (9), we derive the conditions

$$\begin{aligned}
AJ - DI + FH &= 0,\\
BJ - EI + FG &= 0.
\end{aligned} \tag{10}$$

But the equations (6), (9), and (10) show that the Jacobian of our transformation is zero. This proves that $C = 0$.

Since $C = 0$, the equations (7) may be written in the form

$$\begin{aligned}
&-A\bar{p} + B\bar{q} = 0,\\
&D\bar{p} - E\bar{q} = 0,\\
&H\bar{p} - G\bar{q} = 0,\\
&B\bar{y} - E\bar{z} + F\bar{p} = 0,\\
&-A\bar{x} + H\bar{z} - I\bar{q} = 0,\\
&F\bar{x} - I\bar{y} + J\bar{z} = 0,\\
&-B\bar{x} + A\bar{y} + (G - D)\bar{z} - I\bar{p} + F\bar{q} = 0,\\
&E\bar{x} + (-G + qA)\bar{y} - qD\bar{z} + J\bar{p} + qF\bar{q} = 0,\\
&D\bar{x} + (-H - pA)\bar{y} + pD\bar{z} + (J - pF)\bar{q} = 0.
\end{aligned} \tag{11}$$

We shall prove that $\bar{p} = \bar{q} = 0$. Let us suppose that at least one of the quantities $\bar{p}$ and $\bar{q}$ is different from zero. Then from the first three equations of (11), we find

$$\begin{aligned} -AE + BD &= 0, \\ -AG + BH &= 0, \\ -DG + EH &= 0. \end{aligned} \tag{12}$$

In all cases, we shall prove that

$$A\bar{y} - D\bar{z} + F\bar{q} = 0. \tag{13}$$

Let us consider the case in which

$$B = E = G = 0.$$

In this event, it is seen that $\bar{p} = 0$. For if $\bar{p} \neq 0$, then $A = D = H = 0$. Hence, since $C = 0$, it follows by (6) that the Jacobian of the transformation is zero. This proves that $\bar{p} = 0$. Then in this case, the eighth of equations (11) gives equation (13).

Let us now consider the case in which at least one of the quantities B, E, G is not zero. Now by equations (11) and (12), it is seen that

$$\begin{aligned} B(A\bar{y} - D\bar{z} + F\bar{q}) &= 0, \\ E(A\bar{y} - D\bar{z} + F\bar{q}) &= 0, \\ G(A\bar{y} - D\bar{z} + F\bar{q}) &= 0. \end{aligned} \tag{14}$$

Hence since one of the quantities B, E, G is not zero, it follows from (14) that we obtain the equation (13). Thus in all cases we have the equation (13).

From (11) and (13) we then obtain

$$\begin{aligned} -A\bar{p} + B\bar{q} &= 0, \\ D\bar{p} - E\bar{q} &= 0, \\ H\bar{p} - G\bar{q} &= 0, \\ B\bar{y} - E\bar{z} + F\bar{p} &= 0, \\ -A\bar{x} + H\bar{z} - I\bar{q} &= 0, \\ F\bar{x} - I\bar{y} + J\bar{z} &= 0, \\ -B\bar{x} + G\bar{z} - I\bar{p} &= 0, \\ E\bar{x} - G\bar{y} + J\bar{p} &= 0, \\ D\bar{x} - H\bar{y} + J\bar{q} &= 0, \\ A\bar{y} - D\bar{z} + F\bar{q} &= 0. \end{aligned} \tag{15}$$

From (15) it is observed that not all the quantities $\bar{x}$, $\bar{y}$, $\bar{z}$ can be zero. For then since at least one of the quantities $\bar{p}$, $\bar{q}$ is not zero, we must have

$$F = I = J = 0. \tag{16}$$

From $C = 0$, (6), (12) and (16), it follows that the Jacobian of the transformation is zero. Hence at least one of the quantities $\bar{x}$, $\bar{y}$, $\bar{z}$ is not zero.

From (15) we find

$$\begin{aligned} &(AJ - DI + FH)\bar{x} = 0, \qquad (BJ - EI + FG)\bar{x} = 0, \\ &(AJ - DI + FH)\bar{y} = 0, \qquad (BJ - EI + FG)\bar{y} = 0, \\ &(AJ - DI + FH)\bar{z} = 0, \qquad (BJ - EI + FG)\bar{z} = 0. \end{aligned} \tag{17}$$

Then, since at least one of the quantities $\bar{x}$, $\bar{y}$, $\bar{z}$ is not zero, it follows from (17) that

$$\begin{aligned} AJ - DI + FH &= 0, \\ BJ - EI + FG &= 0. \end{aligned} \tag{18}$$

Then since $C = 0$, it follows from (6), (12) and (18) that the Jacobian of the transformation is zero. This proves in all cases that $\bar{p} = \bar{q} = 0$.

Since $\bar{p} = \bar{q} = 0$, the equations (11) now take the form

$$\begin{aligned} B\bar{y} - E\bar{z} &= 0, \\ -A\bar{x} + H\bar{z} &= 0, \\ F\bar{x} - I\bar{y} + J\bar{z} &= 0, \\ -B\bar{x} + A\bar{y} + (G - D)\bar{z} &= 0, \\ E\bar{x} + (-G + qA)\bar{y} - qD\bar{z} &= 0, \\ D\bar{x} + (-H - pA)\bar{y} + pD\bar{z} &= 0. \end{aligned} \tag{19}$$

Let

$$\bar{x} + p\bar{z} = \lambda, \qquad \bar{y} + q\bar{z} = \mu. \tag{20}$$

Then the equations (19) assume the form

$$\begin{aligned} (qB + E)\bar{z} - B\mu &= 0, \\ (pA + H)\bar{z} - A\lambda &= 0, \\ (-pF + qI + J)\bar{z} + F\lambda - I\mu &= 0, \\ (pB - qA + G - D)\bar{z} - B\lambda + A\mu &= 0, \\ (qB + E)\lambda - (pB + G)\mu &= 0, \\ (qA + D)\lambda - (pA + H)\mu &= 0. \end{aligned} \tag{21}$$

At least one of the quantities $\bar{z}$, λ, μ is not zero. For otherwise $\bar{z} = \lambda = \mu = 0$, and hence $\bar{x} = \bar{y} = \bar{z} = \bar{p} = \bar{q} = 0$. This is impossible.

We shall prove that $\bar{z} \neq 0$. Let us suppose that $\bar{z} = 0$. Then equations (21) assume the form

$$
\begin{aligned}
&B\mu = 0, \qquad A\lambda = 0,\\
&-B\lambda + A\mu = 0, \qquad F\lambda - I\mu = 0,\\
&(qB + E)\lambda - (pB + G)\mu = 0,\\
&(qA + D)\lambda - (pA + H)\mu = 0.
\end{aligned} \tag{22}
$$

Since $\bar{z} = 0$, we must have at least one of the quantities λ, μ different from zero. The conditions that equations (22) have a solution for λ, μ at least one of which is different from zero are

$$
\begin{aligned}
&A = 0, \qquad B = 0,\\
&-DG + EH = 0,\\
&-DI + FH = 0,\\
&-EI + FG = 0.
\end{aligned} \tag{23}
$$

From (6) and (23) and the fact that $C = 0$, we see that the Jacobian of the transformation is zero. Hence $\bar{z} \neq 0$.

We shall prove that $qB + E = 0$. Let us suppose that this is not the case. Then we have from (21) that

$$
\bar{z} = \frac{B\mu}{qB + E}, \qquad \lambda = \frac{(pB + G)\mu}{qB + E}. \tag{24}
$$

Since $\bar{z} \neq 0$, we see that $B \neq 0$ and $\mu \neq 0$. Upon substituting (24) into the remaining equations (21), we obtain the conditions

$$
\begin{aligned}
&-AG + BH = 0,\\
&BJ - EI + FG = 0,\\
&-AE + BD = 0,\\
&-DG + EH = 0.
\end{aligned} \tag{25}
$$

From (25) we obtain, since $B \neq 0$,

$$
AJ - DI + FH = 0. \tag{26}
$$

Then from $C = 0$, (6), (25) and (26), we see that the Jacobian of the transformation vanishes. Thus $qB + E = 0$.

We shall prove that $pA + H = 0$. Let us suppose that this is not the case. Then from (21), we obtain

$$
\bar{z} = \frac{A\lambda}{pA + H}, \qquad \mu = \frac{(qA + D)\lambda}{pA + H}. \tag{27}
$$

Since $\bar{z} \gtrless 0$, then A and λ are each different from zero. Substituting (27) into (21), we obtain

$$\begin{aligned} -AE + BD &= 0, \\ AJ - DI + FH &= 0, \\ -AG + BH &= 0, \\ -DG + EH &= 0. \end{aligned} \tag{28}$$

From (28) we obtain, since $A \gtrless 0$,

$$BJ - EI + FG = 0. \tag{29}$$

Then from $C = 0$, (6), (28) and (29), we find that the Jacobian of the transformation is zero. This proves that $pA + H = 0$.

Since $qB + E = 0$ and $pA + H = 0$, our equations (21) become

$$\begin{aligned} B\mu = 0, \quad & A\lambda = 0, \\ G\mu = 0, \quad & D\lambda = 0, \\ (-pF + qI + J)\bar{z} + F\lambda - I\mu &= 0, \\ (pB - qA + G - D)\bar{z} - B\lambda + A\mu &= 0. \end{aligned} \tag{30}$$

We shall prove that $\lambda = \mu = 0$. Let us first suppose that $\lambda \gtrless 0$, $\mu \gtrless 0$. Then we have $A = B = D = E = G = H = 0$. This makes the Jacobian zero. Hence we cannot have $\lambda \gtrless 0$, $\mu \gtrless 0$.

In the second place, let us suppose that $\lambda = 0$, $\mu \gtrless 0$. Then we have

$$\begin{aligned} B = E = G &= 0, \\ (-pF + qI + J)\bar{z} - I\mu &= 0, \\ (-qA - D)\bar{z} + A\mu &= 0. \end{aligned} \tag{31}$$

Hence since $\bar{z} \gtrless 0$, we find from (31) that

$$\begin{aligned} B = E = G &= 0, \\ AJ - DI + FH &= 0. \end{aligned} \tag{32}$$

From $C = 0$, (6) and (32), we see that the Jacobian of the transformation is zero. This gives a contradiction so that we cannot have $\lambda = 0$, $\mu \gtrless 0$.

Finally let us consider the case $\lambda \gtrless 0$, $\mu = 0$. Then our equations (30) become

$$\begin{aligned} A = D = H &= 0, \\ (-pF + qI + J)\bar{z} + F\lambda &= 0, \\ (pB + G)\bar{z} - B\lambda &= 0. \end{aligned} \tag{33}$$

Hence from (33) we obtain

$$A = D = H = 0,$$
$$BJ - EI + FG = 0. \tag{34}$$

This makes the Jacobian of the transformation zero. Hence we cannot have $\lambda \neq 0, \mu = 0$.

Thus by (4) the only possible solutions of our equations (7) are

$$\begin{aligned} Z_x + pZ_z &= P(X_x + pX_z) + Q(Y_x + pY_z), \\ Z_y + qZ_z &= P(X_y + qX_z) + Q(Y_y + qY_z), \\ Z_p &= PX_p + QY_p, \\ Z_q &= PX_q + QY_q, \end{aligned} \tag{35}$$

where

$$\begin{aligned} C = 0, \quad qB + E = 0, \quad pA + H = 0, \\ -pF + qI + J = 0, \\ pB - qA + G - D = 0. \end{aligned} \tag{36}$$

By differentiation the equations (36) are easily found to be consequences of the equations (35), which are identical with Lie's equations (L). Also the equations (35) together with (36) are easily seen to satisfy the equations (7) identically. Thus the validity of our theorem is finally established.

4. **The contact transformations of the infinite group of §3.** It is easy to prove that *the only contact transformations in our group of lineal element transformations are the conformal transformations.*

For the only lineal element contact transformations in space are the extended point transformations (Theorem of Lie). Hence imposing the conditions that our transformation be an extended point transformation, we obtain the conditions

$$\begin{aligned} X_x^2 + Y_x^2 + Z_x^2 = X_y^2 + Y_y^2 + Z_y^2 = X_z^2 + Y_z^2 + Z_z^2, \\ X_yX_z + Y_yY_z + Z_yZ_z = X_zX_x + Y_zY_x + Z_zZ_x \\ = X_xX_y + Y_xY_y + Z_xZ_y = 0. \end{aligned} \tag{37}$$

These are the conditions necessary and sufficient for a conformal transformation; hence our assertion is established. According to Liouville's theorem, the conformal group in space is merely the ten-parameter inversion group.

5. **General transformations.** An element transformation of surface elements, which is not a contact transformation, will convert in general *some* particular integrable systems into integrable systems. That is, a lineal element transformation, not of our group, will convert *some* normal congruences into normal congruences. This subject will be discussed elsewhere.

COLUMBIA UNIVERSITY.

THE SUBGROUP OF ORDER n OF A TRANSITIVE GROUP OF DEGREE n AND CLASS $n - 1$

By Louis Weisner

It is known that a transitive permutation group of degree n and class $n - 1$ has an invariant subgroup of order n consisting of the identity and $n - 1$ permutations of degree n.[1] The object of this paper is to demonstrate that this subgroup of order n is an *Abelian* group.[2]

A permutation of degree $n - 1$ of a transitive permutation group of degree n and class $n - 1$ generates, with the unique subgroup N of order n, a group G which is also a transitive permutation group of degree n and class $n - 1$. We shall confine our attention to this group G. If the order of G is mn, a subgroup M of G that leaves one symbol fixed is a cyclic group of order m,[3] and $G = \{M, N\}$. Let G' be an abstract group simply isomorphic with G, and let M' and N' be the subgroups of G' that correspond to M and N, respectively.

Since N is the commutator subgroup of G, G' has exactly m distinct representations in one variable. Denote these by $\Gamma_1, \cdots, \Gamma_m$, where Γ_1, as usual, denotes the identical representation. Let $\Gamma_{m+1}, \cdots, \Gamma_r$ be the other distinct irreducible representations of G', and let n_ν be the number of variables operated on by Γ_ν.

The relation among group characters

$$\sum_{\nu=1}^{r} \chi_i^{(\nu)} \chi_{i'}^{(\nu)} = g/h_i,$$

where h_i is the number of elements in the i-th conjugate set of G and g is the order of G, becomes, for an element $A \neq 1$ of M',

$$\sum_{\nu=1}^{r} \chi^{(\nu)}(A)\chi^{(\nu)}(A^{-1}) = m.$$

The terms which arise from $\nu = 1, \cdots, m$ have the value 1; hence

$$\chi^{(\nu)}(A) = 0 \qquad (\nu = m + 1, \cdots, r). \tag{1}$$

Received August 29, 1938; presented to the American Mathematical Society, September 6, 1938.

[1] G. Frobenius, *Ueber auflösbare Gruppen* IV, Sitzungsberichte Berlin, 1901, pp. 1223-1225; A. Speiser, *Theorie der Gruppen von endlicher Ordnung*, 3d edition, 1937, p. 202.

[2] The theorem has been proved for the case in which the subgroup that leaves one symbol fixed is of even order. See W. Burnside, *Theory of Groups of Finite Order*, 2d edition, 1911, p. 172.

[3] We could take m to be a prime number, but no advantage is to be derived therefrom.

Therefore

$$\sum_{S \in M'} \chi^{(v)}(S) = n_v \qquad (v = m+1, \cdots, r), \tag{2}$$

and M' has exactly n_v/m linearly independent linear invariants in the representation Γ_v $(v > m)$.[4] Since each of the representations of G' in one variable gives an irreducible representation of M', M' has exactly one independent linear invariant in Γ_1 and no linear invariant in $\Gamma_2, \cdots, \Gamma_m$. Therefore the complete reduction of the permutation group G is given by[5]

$$G = \Gamma_1 + \sum_{v=m+1}^{r} (n_v/m)\Gamma_v. \tag{3}$$

We proceed to prove that $n_v = m$ for each $v > m$.

In the relation

$$\sum_{S \in G'} \chi(S)\bar{\chi}(S) = nm,$$

applied to any Γ_v $(v > m)$, the terms that arise from elements of G' that are not contained in N' vanish by (1). This relation may therefore be written

$$\sum_{S \in N'} \chi(S)\bar{\chi}(S) = nm. \tag{4}$$

Let $\Delta_1, \cdots, \Delta_\rho$ be the distinct irreducible representations of N', Δ_1 being the identical representation. If the reduced form of N' in Γ_v is

$$N_v \equiv \sum_{k=1}^{\rho} a_{vk}\Delta_k, \tag{5}$$

and $\psi^{(k)}(S)$ is the character of S in Δ_k, then

$$\chi(S) = \sum_{k=1}^{\rho} a_{vk}\psi^{(k)}(S).$$

From the relations among group characters and (4), we have

$$m = \sum_{k=1}^{\rho} a_{vk}^2 \qquad (v > m). \tag{6}$$

By (3) and (5) the complete reduction of N is

$$N = \Delta_1 + \sum_{k=1}^{\rho} \sum_{v=m+1}^{r} (n_v/m)a_{vk}\Delta_k. \tag{7}$$

Since N is a regular permutation group, the number of times that Δ_k occurs in the reduction of N is ν_k, the number of variables on which Δ_k operates. Therefore

$$\nu_k = \sum_{v=m+1}^{r} n_v a_{vk}/m \qquad (k > 1). \tag{8}$$

[4] Miller, Blichfeldt and Dickson, *Theory and Applications of Finite Groups*, 1916, p. 258.

[5] Burnside, p. 275.

Moreover, Δ_1 occurs only once. Therefore

$$a_{v1} = 0 \qquad (v > m),$$

and (6) may be written

$$m = \sum_{k=2}^{\rho} a_{vk}^2 \qquad (v > m). \tag{9}$$

Now

$$n = 1 + \sum_{k=2}^{\rho} \nu_k^2.$$

Therefore by (8) and (9)

$$n = 1 + \sum_{k=2}^{\rho} \sum_{v=m+1}^{r} n_v^2 a_{vk}^2/m^2 + 2\sum_{k=2}^{\rho} \sum_{\substack{v,w=m+1 \\ v<w}}^{r} n_v n_w a_{vk} a_{wk}/m^2$$

$$= 1 + \sum_{v=m+1}^{r} n_v^2/m + 2\sum_{k=2}^{\rho} \sum_{\substack{v,w=m+1 \\ v<w}}^{r} n_v n_w a_{vk} a_{wk}/m^2.$$

But, by (3),

$$n = 1 + \sum_{v=m+1}^{r} n_v^2/m.$$

It follows that

$$a_{vk} a_{wk} = 0 \qquad (v \neq w; v > m; w > m).$$

This means that if Δ_k occurs in the reduction of N_v, it cannot occur in the reduction of N_w ($v \neq w$). Therefore in (5)

$$a_{vk} = \nu_k \text{ or } 0.$$

If we suppose that when N_v is reduced $\Delta_2, \cdots, \Delta_s$ are the only irreducible representations that occur, (5) becomes

$$N_v = \sum_{k=2}^{s} \nu_k \Delta_k, \tag{10}$$

and (9) becomes

$$m = \sum_{k=2}^{s} \nu_k^2.$$

Since Δ_k operates on ν_k variables and N_v on n_v variables, we have by (10)

$$n_v = \sum_{k=2}^{s} \nu_k^2.$$

Therefore $n_v = m$, and (3) becomes

$$G = \Gamma_1 + \sum_{v=m+1} \Gamma_v.$$

Hence

$$n = 1 + m(r - m),$$

and

$$r = m + (n - 1)/m$$

is the number of conjugate sets of G.

Of these m are accounted for by the identity and the permutations of degree $n - 1$ of G. Therefore the permutations of degree n of G comprise $(n - 1)/m$ conjugate sets. Since no permutation of degree n of G is commutative with a permutation of degree $n - 1$, the number of conjugate sets into which the permutations of degree n fall with respect to N is m times the number into which they fall with respect to G. Therefore the number of conjugate sets of N, including the identity, is n, and N is an Abelian group.

HUNTER COLLEGE OF THE CITY OF NEW YORK.

IDENTITIES ANALOGOUS TO RAMANUJAN'S IDENTITIES INVOLVING THE PARTITION FUNCTION

By Herbert S. Zuckerman

1. **Introduction.**[1] A new proof of the two identities of Ramanujan

$$\sum_{n=0}^{\infty} p(5n+4)x^n = 5\frac{\varphi(x^5)^5}{\varphi(x)^6}, \tag{1.11}$$

$$\sum_{n=0}^{\infty} p(7n+5)x^n = 7\frac{\varphi(x^7)^3}{\varphi(x)^4} + 49x\frac{\varphi(x^7)^7}{\varphi(x)^8}, \tag{1.12}$$

where $p(n)$ is the number of unrestricted partitions of n and $\varphi(x) = \prod_{n=1}^{\infty}(1-x^n)$, has recently been given in a paper by Rademacher and the author.[2] The method used is not limited to verifying known identities, but can be extended to obtain new ones. In this paper the following new identities will be proved:

$$\begin{aligned}\sum_{n=0}^{\infty} p(25n+24)x^n = {} & 63\cdot 5^2\frac{\varphi(x^5)^6}{\varphi(x)^7} + 52\cdot 5^5 x\frac{\varphi(x^5)^{12}}{\varphi(x)^{13}} + 63\cdot 5^7 x^2\frac{\varphi(x^5)^{18}}{\varphi(x)^{19}} \\ & + 6\cdot 5^{10}x^3\frac{\varphi(x^5)^{24}}{\varphi(x)^{25}} + 5^{12}x^4\frac{\varphi(x^5)^{30}}{\varphi(x)^{31}},\end{aligned} \tag{1.13}$$

$$\begin{aligned}\sum_{n=0}^{\infty} p(49n+47)x^n = {} & 2546\cdot 7^2\frac{\varphi(x^7)^4}{\varphi(x)^5} + 48934\cdot 7^4 x\frac{\varphi(x^7)^8}{\varphi(x)^9} + 1418989\cdot 7^5 x^2\frac{\varphi(x^7)^{12}}{\varphi(x)^{13}} \\ & + 2488800\cdot 7^7 x^3\frac{\varphi(x^7)^{16}}{\varphi(x)^{17}} + 2394438\cdot 7^9 x^4\frac{\varphi(x^7)^{20}}{\varphi(x)^{21}} \\ & + 1437047\cdot 7^{11}x^5\frac{\varphi(x^7)^{24}}{\varphi(x)^{25}} + 4043313\cdot 7^{12}x^6\frac{\varphi(x^7)^{28}}{\varphi(x)^{29}} \\ & + 161744\cdot 7^{15}x^7\frac{\varphi(x^7)^{32}}{\varphi(x)^{33}} + 32136\cdot 7^{17}x^8\frac{\varphi(x^7)^{36}}{\varphi(x)^{37}} \\ & + 31734\cdot 7^{18}x^9\frac{\varphi(x^7)^{40}}{\varphi(x)^{41}} + 3120\cdot 7^{20}x^{10}\frac{\varphi(x^7)^{44}}{\varphi(x)^{45}} \\ & + 204\cdot 7^{22}x^{11}\frac{\varphi(x^7)^{48}}{\varphi(x)^{49}} + 8\cdot 7^{24}x^{12}\frac{\varphi(x^7)^{52}}{\varphi(x)^{53}} + 7^{25}x^{13}\frac{\varphi(x^7)^{56}}{\varphi(x)^{57}},\end{aligned} \tag{1.14}$$

Received September 10, 1938; presented to the American Mathematical Society, April 16, 1938. The author is a National Research Fellow.

[1] Many of the results of this paper have been obtained by G. N. Watson in a paper (Journal für Mathematik, vol. 179(1938), pp. 97–128) which appeared since this was written. The methods of the two papers are quite different.

[2] *A new proof of two of Ramanujan's identities*, to be published in the Annals of Mathematics.

$$\sum_{n=0}^{\infty} p(13n+6)x^n = 11\,\frac{\varphi(x^{13})}{\varphi(x)^2} + 36\cdot 13x\,\frac{\varphi(x^{13})^3}{\varphi(x)^4} + 38\cdot 13^2x^2\,\frac{\varphi(x^{13})^5}{\varphi(x)^6}$$

$$+ 20\cdot 13^3x^3\,\frac{\varphi(x^{13})^7}{\varphi(x)^8} + 6\cdot 13^4x^4\,\frac{\varphi(x^{13})^9}{\varphi(x)^{10}} \tag{1.15}$$

$$+ 13^5x^5\,\frac{\varphi(x^{13})^{11}}{\varphi(x)^{12}} + 13^6x^6\,\frac{\varphi(x^{13})^{13}}{\varphi(x)^{14}}.$$

It is probable that further identities involving higher powers of 5, 7, and 13 as moduli could also be found, but the computational difficulties would be considerable.

From (1.11) and (1.12) Ramanujan deduced the divisibility properties

$$p(25n+24) \equiv 0 \pmod{25}, \tag{1.21}$$

$$p(49n+47) \equiv p(49n+19) \equiv p(49n+33) \equiv p(49n+40) \equiv 0 \pmod{49}. \tag{1.22}$$

In a similar way we shall use (1.13) and (1.14) to find

$$p(125n+99) \equiv p(125n+74) \equiv p(125n+124) \equiv 0 \pmod{125}, \tag{1.23}$$

$$p(625n+599) \equiv 0 \pmod{625}, \tag{1.24}$$

$$p(2401n+2301) \equiv p(2401n+929) \equiv p(2401n+1615)$$

$$\equiv p(2401n+1958) \equiv 0 \pmod{343}. \tag{1.25}$$

The first congruence of (1.23) has already been proved by Krezmer,[3] and it and (1.24) are special cases of Ramanujan's conjecture that $p(Rn+\lambda)$ is divisible by R if $R = 5^a7^b11^c$ and $24\lambda \equiv 1 \pmod{R}$. This conjecture is known to be incorrect for the modulus $R = 343$, but (1.25) can be thought of as a valid substitute.

In the proof of the identities we make use of the modular form

$$\eta(\tau) = e^{\frac{1}{24}\pi i\tau}\prod_{n=1}^{\infty}(1-e^{2\pi i n\tau}). \tag{1.3}$$

Thus, using this function, we may write (1.13) as

$$e^{\frac{1}{24}\pi i\tau}\sum_{n=0}^{\infty} p(25n+24)e^{2\pi i n\tau} = 63\cdot 5^2\,\frac{\eta(5\tau)^6}{\eta(\tau)^7} + 52\cdot 5^6\,\frac{\eta(5\tau)^{12}}{\eta(\tau)^{13}}$$

$$+ 63\cdot 5^7\,\frac{\eta(5\tau)^{18}}{\eta(\tau)^{19}} + 6\cdot 5^{10}\,\frac{\eta(5\tau)^{24}}{\eta(\tau)^{25}} + 5^{12}\,\frac{\eta(5\tau)^{30}}{\eta(\tau)^{31}}, \tag{1.4}$$

and it is in this form that it will be obtained. The other two identities can be expressed in a similar manner.

[3] Bulletin of the Academy of Sciences, U.R.S.S., no. 6(1933), pp. 763–800.

2. **Outline of the method.** We wish to find identities in which one member is of the form

$$e^{2\pi i s r} \sum_{n=0}^{\infty} p(Rn + S) e^{2\pi i n r}. \tag{2.1}$$

Since the functions $p(n)$ are coefficients in the expansion of a modular form of dimension $\frac{1}{2}$, we shall try to build up the other member from modular forms which also have the dimension $\frac{1}{2}$. For these forms we choose the functions

$$F_\rho(\tau) = \frac{\eta(Q\tau)^\rho}{\eta(\tau)^{\rho+1}}, \tag{2.2}$$

where ρ is a positive integer and Q is an odd prime.

In the following section we shall see that $F_\rho(\tau)$ has an expansion of the form

$$F_\rho(\tau) = e^{2\pi i \alpha_\rho r} \sum_{n=0}^{\infty} a_{n,\rho} e^{2\pi i n r}, \tag{2.31}$$

where

$$a_{n,\rho} = \frac{2\pi}{(n + \alpha_\rho)^{\frac{1}{2}} Q^{\frac{1}{2}}} \sum_{\nu=1}^{\mu_\rho} b_{-\nu,\rho} (\nu - \beta_\rho)^{\frac{1}{2}} \cdot \sum_{\substack{k \\ (k,Q)=1}} \frac{1}{k} A_{k,\rho}^{(\nu)}(n) I_{\frac{1}{2}}\left(\frac{4\pi}{k} \sqrt{\frac{(\nu - \beta_\rho)}{Q}(n + \alpha_\rho)}\right),$$

$$\alpha_\rho = \frac{(Q-1)\rho - 1}{24} - \left[\frac{(Q-1)\rho - 1}{24}\right], \tag{2.32}$$

$$\beta_\rho = \left[\frac{(Q-1)\rho + Q}{24}\right] + 1 - \frac{(Q-1)\rho + Q}{24},$$

$$\mu_\rho = \left[\frac{(Q-1)\rho + Q}{24}\right] + 1,$$

and where $\sum_{\substack{k \\ (k,Q)=1}}$ designates the sum over all $k > 0$ such that $(k, Q) = 1$. This value of $a_{n,\rho}$ is to be compared with the known value[4]

$$p(Rn + S) = \frac{2\pi}{(24Rn + 24S - 1)^{\frac{1}{2}}} \cdot \sum_{k=1}^{\infty} \frac{1}{k} A_k(Rn + S) I_{\frac{1}{2}}\left(\frac{4\pi}{k} \sqrt{\frac{1}{24}\left(Rn + S - \frac{1}{24}\right)}\right). \tag{2.4}$$

[4] For this form of the series for $p(n)$ see Rademacher and Zuckerman, *On the Fourier coefficients of certain modular forms of positive dimension*, Annals of Mathematics, vol. 39 (1938), pp. 433–462, in particular, pp. 455–456.

Comparing the arguments of the Bessel functions, we find, in the first place,

$$\frac{24S-1}{24R} = \alpha_\rho,$$

and hence, since α_ρ is an integer divided by 24, we have

$$24S \equiv 1 \pmod{R}. \tag{2.5}$$

Besides this, in order to get the terms of (2.4) in which Q does not divide k, we must have, for some integer $\nu_1 \leqq \mu_\rho$ and some ρ,

$$\frac{\nu_1 - \beta_\rho}{Q} = \frac{R}{24}. \tag{2.61}$$

To get the terms in which Q divides k but Q^2 does not we write the argument in (2.32) for $\nu = \nu_2 \leqq \mu_\rho$ as

$$\frac{4\pi}{Qk}\sqrt{Q(\nu_2 - \beta_\rho)(n+\alpha_\rho)}$$

and demand

$$Q(\nu_2 - \beta_\rho) = \frac{R}{24}. \tag{2.62}$$

Since β_ρ is an integer divided by 24, we see that R is a multiple of Q and therefore, by (2.5), $24S \equiv 1 \pmod{Q}$. Now it is known[5] that $A_{Q^2k}(\lambda) = 0$ if $24\lambda \equiv 1 \pmod{Q}$ and hence no terms for which Q^2 divides k appear in (2.4).

If we are to find an identity involving a single $F_\rho(\tau)$, we see from the above remarks that we must have $\mu_\rho = 2$ and hence $\nu_1 = 2$, $\nu_2 = 1$. Then (2.61) and (2.62) yield the result

$$\frac{2-\beta_\rho}{Q} = Q(1-\beta_\rho), \qquad \beta_\rho = \frac{Q^2-2}{Q^2-1}.$$

Since $Q^2 - 2$ and $Q^2 - 1$ are relatively prime and β_ρ is an integer divided by 24, we have

$$Q^2 - 1 \mid 24,$$

but the only odd primes for which this is satisfied are $Q = 3$ and $Q = 5$. The value $Q = 3$ is clearly impossible since we saw that Q divides R and (2.5) cannot be solved for S if 3 divides R. The case $Q = 5$ is possible and leads to Ramanujan's identity (1.11).

We must therefore use linear combinations of the $F_\rho(\tau)$ in order to find new identities. For the various terms to combine correctly we must choose values of

[5] Loc. cit., footnote 1, especially Lemma 2.

ρ for which α_ρ takes the same value. Hence if ρ_0 is one value of ρ, we must have

$$\frac{(Q-1)\rho-1}{24}-\left[\frac{(Q-1)\rho-1}{24}\right]=\frac{(Q-1)\rho_0-1}{24}-\left[\frac{(Q-1)\rho_0-1}{24}\right],$$

$$(Q-1)\rho-1\equiv(Q-1)\rho_0-1 \pmod{24},$$

$$\rho\equiv\rho_0\left(\mathrm{mod}\,\frac{24}{g}\right),$$

$$\rho=\rho_0+\frac{24}{g}\rho', \tag{2.71}$$

where

$$g=(Q-1,\,24).$$

With these values for ρ we now have

$$\begin{aligned}\mu_\rho&=\left[\frac{(Q-1)\left(\rho_0+\frac{24}{g}\rho'\right)+Q}{24}\right]+1\\&=\left[\frac{(Q-1)\rho_0+Q}{24}\right]+\frac{Q-1}{g}\rho'+1\\&=\mu_{\rho_0}+\frac{Q-1}{g}\rho'.\end{aligned}$$

This means that as ρ increases to its next larger value, μ_ρ increases by $(Q-1)/g$. However, we must choose a linear combination of the $F_\rho(\tau)$ in such a way as to eliminate all the terms in the series except those corresponding to $\nu=\nu_1$ and $\nu=\nu_2$. This can be done only if the μ_ρ increase by unity at each increase of ρ since otherwise there will not be enough functions to allow us to make the elimination.[6] Hence we must have

$$(Q-1,\,24)=g=Q-1,$$

$$Q-1\,|\,24. \tag{2.72}$$

The only permissible values are then $Q=3,\,5,\,7,\,13$, but $Q=3$ is impossible as was shown above.

We have now found that we are to try to find identities by using linear combinations of the $F_\rho(\tau)$ and that we are to use the values $Q=5,\,7,\,13$. The conditions which we have discussed are only necessary ones, and it will remain to show that the α_ρ and β_ρ can be properly determined simultaneously and that the eliminations can actually be accomplished, and to treat the $A_{k,\rho}^{(\nu)}(n)$.

[6] The case $\rho=\rho_0,\,\rho_0+24/g$, $\mu_\rho=1$ and 3, respectively, is not possible since it requires $\nu_1=3$, $\nu_2=2$ and hence, by (2.61) and (2.62), $\beta_\rho>1$.

3. **The expansion of the $F_\rho(\tau)$.** Since the $F_\rho(\tau)$ are modular forms belonging to a subgroup of the modular group, they are a special case of a previous paper.[7] However, we can obtain their expansions in a more convenient form directly. Except for one change the proof follows that used by Rademacher and the author[8] in determining the expansions of forms belonging to the full modular group.

From (1.3) and (2.2) we have

$$F_\rho(\tau) = \left[e^{\frac{1}{12}\pi i r Q\rho}\prod_{n=1}^{\infty}(1-e^{2\pi i Q n\tau})^\rho\right]\Big/\left[e^{\frac{1}{12}\pi i r(\rho+1)}\prod_{n=1}^{\infty}(1-e^{2\pi i n\tau})^{\rho+1}\right] \tag{3.11}$$

$$= e^{2\pi i a_{\rho,r}}\sum_{n=0}^{\infty}a_{n,\rho}e^{2\pi i n\tau} = e^{2\pi i a_{\rho,r}}f_\rho(e^{2\pi i\tau}),$$

where

$$\alpha_\rho = \frac{(Q-1)\rho-1}{24} - \left[\frac{(Q-1)\rho-1}{24}\right], \tag{3.12}$$

and hence $f_\rho(x)$ does not have any terms with negative exponent. The proof now follows that of the other paper except for the following change in §2 of that paper. For k divisible by Q we still use the transformation

$$\begin{pmatrix} a & b \\ c & d\end{pmatrix} = \begin{pmatrix} h^* & -\dfrac{hh^*+1}{k} \\ k & -h\end{pmatrix},$$

where h^* is a solution of $hh^* \equiv -1 \pmod{k}$. Applying this transformation and using the transformation equation

$$\eta\left(\frac{a\tau+b}{c\tau+d}\right) = e^{\frac{1}{12}\pi i(a+d)/c}\omega_{a,c}^{-1}(-i(c\tau+d))^{\frac{1}{2}}\eta(\tau), \qquad c>0, \tag{3.2}$$

we have

$$F_\rho(\tau') = F_\rho\left(\frac{h^*\tau - k^{-1}(hh^*+1)}{k\tau-h}\right)$$

$$= \eta\left(\frac{h^*Q\tau - Qk^{-1}(hh^*+1)}{kQ^{-1}Q\tau-h}\right)^\rho \Big/ \eta\left(\frac{h^*\tau-k^{-1}(hh^*+1)}{k\tau-h}\right)^{\rho+1}$$

$$= \exp\{\tfrac{1}{12}\pi i k^{-1}[(\rho Q^2-\rho-1)h^* - (\rho Q-\rho-1)h]\}\,\frac{\omega_{h^*,k}^{\rho+1}}{\omega_{h^*Q,k/Q}^{\rho}}\,[-i(k\tau-h)]^{-\frac{1}{2}}\,\frac{\eta(Q\tau)^\rho}{\eta(\tau)^{\rho+1}}$$

$$= \exp\{\tfrac{1}{12}\pi i k^{-1}[(\rho Q^2-\rho-1)h^* - (\rho Q-\rho-1)h]\}\,\frac{\omega_{h^*,k}^{\rho+1}}{\omega_{h^*Q,k/Q}^{\rho}}\,[-i(k\tau-h)]^{-\frac{1}{2}}F_\rho(\tau).$$

[7] *On the coefficients of certain modular forms belonging to subgroups of the modular group*, to be published in the Transactions of the American Mathematical Society.

[8] Loc. cit., footnote 3.

For these k the proof now proceeds as in the other paper and, since $f_\rho(x)$ has no terms with negative exponents, §3 shows that these k will not appear in the final result.

For k relatively prime to Q we use the transformation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} h & -\dfrac{Qhh'+1}{k} \\ k & -Qh' \end{pmatrix},$$

where $Qhh' \equiv -1 \pmod{k}$, and have, as before,

$$F_\rho(\tau') = F_\rho\left(\frac{h\tau - \dfrac{Qhh'+1}{k}}{k\tau - Qh'}\right)$$

$$(3.31)\quad = \eta\left(\frac{Qh\tau Q^{-1} - k^{-1}(Qhh'+1)}{k\tau Q^{-1} - h'}\right)^\rho \Big/ \eta\left(\frac{h\tau - k^{-1}(Qhh'+1)}{k\tau - Qh'}\right)^{\rho+1}$$

$$= \exp\{\tfrac{1}{12}\pi i k^{-1}[(Q\rho - \rho - 1)h + (Q\rho + Q - \rho)h']\} \cdot \frac{\omega_{h,k}^{\rho+1}}{\omega_{Qh,k}^{\rho}}\, Q^{-\frac{1}{2}\rho}[-i(k\tau - h')]^{-\frac{1}{2}} \frac{\eta(\tau Q^{-1})^\rho}{\eta(\tau)^{\rho+1}}.$$

Now by (1.3) we have

$$(3.32)\quad Q^{-\frac{1}{2}\rho}\frac{\eta(\tau Q^{-1})^\rho}{\eta(\tau)^{\rho+1}} = Q^{-\frac{1}{2}\rho} e^{\frac{1}{12}\pi i \tau\rho/Q} \prod_{n=1}^{\infty}(1 - e^{2\pi i n\tau/Q})^\rho \Big/ e^{\frac{1}{12}\pi i \tau(\rho+1)} \prod_{n=1}^{\infty}(1 - e^{2\pi i n\tau})^{\rho+1}$$

$$= e^{2\pi i \beta_\rho \tau/Q} \sum_{n=-\mu_\rho}^{\infty} b_{n,\rho} e^{2\pi i n\tau/Q} = e^{2\pi i \beta_\rho \tau/Q} g_\rho(e^{2\pi i n\tau/Q}),$$

where

$$(3.33)\quad \beta_\rho = \left[\frac{(Q-1)\rho + Q}{24}\right] + 1 - \frac{(Q-1)\rho + Q}{24}, \quad \mu_\rho = \left[\frac{(Q-1)\rho + Q}{24}\right] + 1.$$

We now take

$$\tau = \frac{i}{kz} + \frac{Qh'}{k}, \qquad \tau' = \frac{iz}{k} + \frac{h}{k}$$

and combine (3.11), (3.31) and (3.32) to obtain

$$f_\rho(e^{2\pi(z - ih)/k}) = \Omega_{h,k}\Psi_k(z) g_\rho(\exp[2\pi i h' k^{-1} - 2\pi/(Qkz)])$$

with

$$\Omega_{h,k} = \exp\{\tfrac{1}{12}\pi i k^{-1}[(Q\rho - \rho - 1)h + (Q\rho + Q - \rho)h']\} \cdot \frac{\omega_{h,k}^{\rho+1}}{\omega_{Qh,k}^{\rho}} \exp[2\pi i(\beta_\rho h'/k - \alpha_\rho h/k)],$$

$$\Psi_k(z) = z^{\frac{1}{2}} \exp\{2\pi[\alpha_\rho z/k - \beta_\rho/(Qkz)]\}.$$

The proof now continues as before but we must replace ϵ^{-1} by

$$\exp\{\tfrac{1}{12}\pi i k^{-1}[(Q\rho-\rho-1)h+(Q\rho+Q-\rho)h']\}\,\frac{\omega_{h,k}^{\rho+1}}{\omega_{Qh,k}^{\rho}},$$

the $a_{-\nu}$ by $b_{-\nu,\rho}$, τ by $\frac{1}{2}$, h' by Qh', and the $\Omega_{h,k}$ and $\Psi_k(z)$ by the above values. The integrand (4.42) of the other paper now becomes

$$w^{\frac{1}{2}}\exp[2\pi(m+\alpha_\rho)w+(\nu-\beta_\rho)/(Qk^2w)]$$

so that we are to replace $m+\alpha$ by $m+\alpha_\rho$ and $\nu-\alpha$ by $(\nu-\beta_\rho)/Q$ in the final result, (5.5) and (5.6), of that paper. Thus we have

$$a_{m,\rho}=2\pi\sum_{\nu=1}^{\mu_\rho}b_{-\nu,\rho}$$

(3.41)
$$\cdot\sum_{\substack{k\\(k,Q)=1}}\frac{1}{k}A_{k,\rho}^{(\nu)}(m)\left(\frac{\nu-\beta_\rho}{Q(m+\alpha_\rho)}\right)^{\frac{1}{2}}I_1\left(\frac{4\pi}{k}\sqrt{\frac{\nu-\beta_\rho}{Q}(m+\alpha_\rho)}\right),$$

$$A_{k,\rho}^{(\nu)}(m)=\sum_{\substack{h \bmod k\\(h,k)=1}}\exp\{\tfrac{1}{12}\pi i[(Q\rho-\rho-1)h+(Q\rho+Q-\rho)h']/k\}$$

(3.42)
$$\cdot\frac{\omega_{h,k}^{\rho+1}}{\omega_{Qh,k}^{\rho}}\exp\{-2\pi i[(\nu-\beta_\rho)h'+(m+\alpha_\rho)h]/k\},$$

which is the same as (2.32).

4. **The $A_{k,\rho}^{(\nu)}(m)$.** Before we consider the three values of Q separately, it will be convenient to set down certain facts pertaining to the $A_{k,\rho}^{(\nu)}(m)$. By (2.71) and (2.72) we see that in each case we shall have a fixed ρ_0 and shall take all of our ρ of the form

$$\rho=\rho_0+\frac{24}{Q-1}\rho'.$$

We shall now show that for these ρ we have

(4.1)
$$A_{k,\rho}^{(\nu)}(m)=A_{k,\rho_0}^{(\nu)}(m).$$

By (3.42) it is clearly sufficient to prove

$$\exp\left\{\tfrac{1}{12}\pi i k^{-1}\left[\left(\frac{24Q}{Q-1}-\frac{24}{Q-1}\right)h+\left(\frac{24Q}{Q-1}-\frac{24}{Q-1}\right)h'\right]\right\}\frac{\omega_{h,k}^{24/(Q-1)}}{\omega_{Qh,k}^{24/(Q-1)}}=1$$

which reduces to

(4.2)
$$\exp[2\pi i k^{-1}(h+h')]\,\frac{\omega_{h,k}^{24/(Q-1)}}{\omega_{Qh,k}^{24/(Q-1)}}=1.$$

We use the value of $\omega_{h,k}$ as given by the formulas

(4.31)
$$\omega_{h,k}=\left(\frac{-k}{h}\right)\exp\{-[\tfrac{1}{4}(2-hk-h)+\tfrac{1}{12}(k-k^{-1})(2h-h^*+h^2h^*)]\pi i\}$$

for h odd, and

$$\omega_{h,k} = \left(\frac{-h}{k}\right) \exp\{-[\tfrac{1}{4}(k-1) + \tfrac{1}{12}(k-k^{-1})(2h - h^* + h^2h^*)]\pi i\} \tag{4.32}$$

for k odd, where h^* is again any solution of $hh^* \equiv -1 \pmod{k}$. Now, since for $Q = 5, 7$, or 11, the value $24/(Q-1)$ is an even integer, we have, for k even and hence h odd,

$$\begin{aligned}
&\exp[2\pi ik^{-1}(h+h')] \frac{\omega_{h,k}^{24/(Q-1)}}{\omega_{Qh,k}^{24/(Q-1)}} \\
&\quad = \exp[2\pi ik^{-1}(h+h')] \\
&\quad \cdot \exp\{-24\pi i(Q-1)^{-1}[\tfrac{1}{4}(2-hk-h) + \tfrac{1}{12}(k-k^{-1})(2h - Qh' + Qh^2h')]\} \\
&\quad \cdot \exp\{24\pi i(Q-1)^{-1}[\tfrac{1}{4}(2-Qhk-Qh) + \tfrac{1}{12}(k-k^{-1})(2Qh - h' + Q^2h^2h')]\},
\end{aligned}$$

where we have used the definition $Qhh' \equiv -1 \pmod{k}$ to eliminate h^*. Simplifying the above, we find

$$\begin{aligned}
\exp[2\pi ik^{-1}(h+h')] \frac{\omega_{h,k}^{24/(Q-1)}}{\omega_{Qh,k}^{24/(Q-1)}} &= \exp[2\pi i(-hk - 3h + h'k + Qh^2h'k)] \\
&\quad \cdot \exp[-2\pi ik^{-1}(h + Qh^2h')] \\
&= \exp[-2\pi ihk^{-1}(Qhh' + 1)] = 1.
\end{aligned}$$

Similarly if k is odd, we have

$$\begin{aligned}
&\exp[2\pi ik^{-1}(h+h')] \frac{\omega_{h,k}^{24/(Q-1)}}{\omega_{Qh,k}^{24/(Q-1)}} \\
&\quad = \exp[2\pi ik^{-1}(h+h')] \\
&\qquad \cdot \exp\{-24\pi i(Q-1)^{-1}[\tfrac{1}{4}(k-1) + \tfrac{1}{12}(k-k^{-1})(2h - Qh' + Qh^2h')]\} \\
&\qquad \cdot \exp\{24\pi i(Q-1)^{-1}[\tfrac{1}{4}(k-1) + \tfrac{1}{12}(k-k^{-1})(2Qh - h' + Q^2h^2h')]\} \\
&\quad = \exp[2\pi ik(2h + h' + Qh^2h')] \exp[-2\pi ik^{-1}(h + Qh^2h')] \\
&\quad = \exp[-2\pi ihk^{-1}(Qhh' + 1)] = 1.
\end{aligned}$$

For later reference we note several facts concerning the $A_k(n)$ which are given by the formula

$$A_k(n) = \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{h,k} e^{-2\pi inh/k}. \tag{4.41}$$

The values of R in which we shall be interested are $R = Q$ or Q^2 so we shall limit ourselves to these two cases. In §2 we found

$$A_{Q^2k}(Rn + S) = 0. \tag{4.42}$$

In addition to this we shall need expressions for $A_{Qk}(Rn + S)$ and $A_k(Rn + S)$, where k is relatively prime to Q. For $R = Q^2$ we have

$$A_k(Q^2n + S) = \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{h,k} \exp\left[-2\pi i h k^{-1}(Q^2n + S)\right].$$

Now, since $Q^2 - 1$ is a multiple of 24, (4.31) and (4.32) yield, for k relatively prime to Q,

$$\frac{\omega_{Q^2h,k}}{\omega_{h,k}} = \exp\left[\tfrac{1}{12}\pi i k^{-1}(Q^2 - 1)(2h + h'' + Q^2h^2h'')\right], \tag{4.43}$$

where h'' is any solution of $Q^2hh'' \equiv -1 \pmod{k}$. Using this, we have

$$\begin{aligned} A_k(Q^2n + S) &= \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{Q^2h,k} \exp\left\{-2\pi i k^{-1}[(Q^2n + S + \tfrac{1}{24}(Q^2 - 1))h \right. \\ &\qquad \left. + \tfrac{1}{24}(Q^2 - 1)h'']\right\} \exp\left\{-2\pi i \cdot \tfrac{1}{24}(Q^2 - 1)h \cdot (1 + Q^2hh'')k^{-1}\right\} \\ &= \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{Q^2h,k} \exp\left\{-2\pi i k^{-1}[(n + \tfrac{1}{24}Q^{-2}(24S - 1 + Q^2))Q^2h \right. \\ &\qquad \left. + \tfrac{1}{24}(Q^2 - 1)h'']\right\} \qquad (4.44) \\ &= \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{h,k} \exp\left\{-2\pi i k^{-1}[(n + \tfrac{1}{24}Q^{-2}(24S - 1 + Q^2)h \right. \\ &\qquad \left. + \tfrac{1}{24}(Q^2 - 1)Qh']\right\}, \end{aligned}$$

where $Qhh' \equiv -1 \pmod{k}$ and where the last step was accomplished by introducing a new index of summation.

For $R = Q$ we use (4.43) to find

$$\begin{aligned} A_k(Qn + S) &= \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{h,k} \exp\left[-2\pi i k^{-1}(Qn + S)h\right] \\ &= \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{Q^2h,k} \exp\left\{-2\pi i k^{-1}[(Qn + S + \tfrac{1}{24}(Q^2 - 1))h \right. \\ &\qquad \left. + \tfrac{1}{24}(Q^2 - 1)h'']\right\} \exp\left[2\pi i \cdot \tfrac{1}{24}(Q^2 - 1)h \cdot (1 + Q^2hh'')k^{-1}\right] \\ &= \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{Q^2h,k} \exp\left\{-2\pi i k^{-1}[(n + \tfrac{1}{24}Q^{-1}(24S - 1 + Q^2))Qh \right. \\ &\qquad \left. + \tfrac{1}{24}(Q^2 - 1)h'']\right\} \qquad (4.45) \\ &= \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{Qh,k} \exp\left\{-2\pi i k^{-1}[(n + \tfrac{1}{24}Q^{-1}(24S - 1 + Q^2))h \right. \\ &\qquad \left. + \tfrac{1}{24}(Q^2 - 1)h']\right\}. \end{aligned}$$

To obtain expressions for $A_{Qk}(Rn + S)$ for k relatively prime to Q, we use the fact[9] that

$$A_{pk}(pn + \lambda) = A_p(\lambda)A_k(n_1), \tag{4.51}$$

[9] Loc. cit., footnote 1, Lemma 1.

where

$$p^2 n_1 \equiv pn + \tfrac{1}{24}(p^2 - 1 + 24\lambda) \pmod{k} \tag{4.52}$$

if p is a prime > 3, $p \nmid k$ and $24\lambda \equiv 1 \pmod{p}$. For $R = Q$ we may take

$$S = Q - \frac{Q^2 - 1}{24},$$

which, for $Q < 24$, determines S as the least positive solution of $24S \equiv 1 \pmod{Q}$. Then by (4.51) and (4.52) we have

$$A_{Qk}(Qn + S) = A_Q(S) A_k(n_1),$$

where

$$Q^2 n_1 \equiv Qn + Q \pmod{k}.$$

If we now define $\bar{k}$ by the congruence $k\bar{k} \equiv 1 \pmod{Q}$, we have

$$Q^2 n_1 \equiv (Qn + Q)(1 - k\bar{k}) \pmod{k},$$

$$n_1 \equiv (n + 1)\frac{1 - k\bar{k}}{Q} \pmod{k}.$$

Now it is known[10] that

$$A_p(\lambda) = \left(\frac{3}{p}\right)\sqrt{p} \tag{4.53}$$

if $24\lambda \equiv 1 \pmod{p}$, p a prime > 3. Combining these results, we then have the desired result

$$A_{Qk}(Qn + S) = \left(\frac{3}{Q}\right)\sqrt{Q}\, A_k\left((n + 1)\frac{1 - k\bar{k}}{Q}\right) \tag{4.6}$$

provided $S = Q - \frac{1}{24}(Q^2 - 1)$.

For $R = Q^2$ we may take

$$S = \tfrac{1}{24}(23Q^2 + 1),$$

which determines S again as the least positive solution of its defining congruence, $24S \equiv 1 \pmod{Q^2}$. Applying (4.6) we then find

$$\begin{aligned} A_{Qk}(Q^2 n + S) &= A_{Qk}[Q(Qn + Q - 1) + Q - \tfrac{1}{24}(Q^2 - 1)] \\ &= \left(\frac{3}{Q}\right)\sqrt{Q}\, A_k\left[(Qn + Q)\frac{1 - k\bar{k}}{Q}\right] \end{aligned}$$

and therefore, since $A_k(n)$ has the period k,

$$A_{Qk}(Q^2 n + S) = \left(\frac{3}{Q}\right)\sqrt{Q}\, A_k(n + 1). \tag{4.7}$$

[10] D. H. Lehmer, *On the series for the partition function*, Transactions of the American Mathematical Society, vol. 43(1938), pp. 271–295, especially Theorem 5.

5. **The case $Q = 5$.** For $Q = 5$ we take ρ of the form $\rho = \rho_0 + 6\rho'$, where we may choose ρ_0 from the values $\rho_0 = 1, 2, 3, 4, 5, 6$. However, it can be shown that the conditions found in §2 are not satisfied for $\rho_0 = 1$ or 4. Of the remaining values we shall limit ourselves to the values $\rho_0 = 5$ or 6 which lead to the identities (1.11) and (1.13) and possibly others.

We first consider $\rho_0 = 6$ and have

$$\alpha_\rho = \tfrac{23}{24}, \qquad \beta_\rho = \tfrac{19}{24}, \qquad \mu_\rho = \rho' + 2.$$

The conditions of §2 then become

$$\frac{24S - 1}{24R} = \frac{23}{24}, \qquad 24\nu_1 - 19 = 5R, \qquad 24\nu_2 - 19 = \frac{R}{5}$$

which reduce to

$$R \equiv 1 \pmod{24}, \qquad \nu_1 = \frac{5R + 19}{24}, \qquad \nu_2 = \frac{R + 95}{120}.$$

In particular R may be taken as any even power of 5, and it is probable that we could then obtain an identity. Other values might also be taken for R but the later discussion of the $A_{k,\rho}^{(\nu)}(m)$ is more difficult and may not be possible. We take $R = 25$ and then have $\nu_1 = 6$, $\nu_2 = 1$, so we take $\rho' = 0, 1, 2, 3, 4$ and write

$$F(\tau) = \sum_{\rho'=0}^{4} A_{\rho'} F_{6+6\rho'}(\tau) = e^{2\pi i \frac{19}{24}\tau} \sum_{n=0}^{\infty} a_n e^{2\pi i n\tau}. \tag{5.1}$$

Then, by (2.31) and (2.32), we have

$$\begin{aligned} a_n &= \sum_{\rho'=0}^{4} A_{\rho'} a_{n,6+6\rho'} \\ &= \sum_{\rho'=0}^{4} A_{\rho'} \frac{2\pi}{(n + \frac{23}{24})^{\frac{1}{4}} 5^{\frac{1}{4}}} \sum_{\nu=1}^{\rho'+2} b_{-\nu,6+6\rho'} (\nu - \tfrac{19}{24})^{\frac{1}{4}} \\ &\qquad \cdot \sum_{\substack{k \\ (k,5)=1}} \frac{1}{k} A_{k,6+6\rho'}^{(\nu)}(n) I_{\frac{1}{4}}\left(\frac{4\pi}{k}\sqrt{\frac{24\nu - 19}{5 \cdot 24}\,(n + \tfrac{23}{24})}\right) \\ &= \frac{2\pi \cdot 5^{\frac{1}{4}}}{(24 \cdot 25n + 24 \cdot 24 - 1)^{\frac{1}{4}}} \sum_{\nu=1}^{6} (24\nu - 19)^{\frac{1}{4}} \sum_{\nu-2 \le \rho' \le 4} A_{\rho'} b_{-\nu,6+6\rho'} \\ &\qquad \cdot \sum_{\substack{k \\ (k,5)=1}} \frac{1}{k} A_{k,6}^{(\nu)}(n) I_{\frac{1}{4}}\left(\frac{4\pi}{k}\sqrt{\frac{24\nu - 19}{5 \cdot 24}\,(n + \tfrac{23}{24})}\right), \end{aligned}$$

where we have used (4.1) to replace $A_{k,6+6\rho'}^{(\nu)}(n)$ by $A_{k,6}^{(\nu)}(n)$. If we now choose the $A_{\rho'}$ as solutions of

$$A_4 b_{-6,30} = 1, \qquad \sum_{\nu-2 \le \rho' \le 4} A_{\rho'} b_{-\nu,6+6\rho'} = 0 \quad \text{for } \nu = 2, 3, 4, 5, \tag{5.2}$$

we have

$$a_n = \frac{2\pi\cdot 5^{\frac{1}{2}}}{(24\cdot 25n + 24\cdot 24 - 1)^{\frac{1}{4}}}\, 5^{\frac{1}{2}} \sum_{\rho'=0}^{4} A_{\rho'} b_{-1,6+8\rho'} \cdot \sum_{\substack{k \\ (k,5)=1}} \frac{1}{k} A_{k,6}^{(1)}(n) I_{\frac{1}{2}}\left(\frac{4\pi}{k}\sqrt{\frac{1}{24}\left(n+\frac{23}{24}\right)}\right)$$

$$+ \frac{2\pi\cdot 5^{\frac{1}{2}}}{(24\cdot 25n + 24\cdot 24 - 1)^{\frac{1}{4}}}\, 125^{\frac{1}{4}} \sum_{\substack{k \\ (k,5)=1}} \frac{1}{k} A_{k,6}^{(6)}(n) I_{\frac{1}{2}}\left(\frac{4\pi}{k}\sqrt{\frac{25}{24}\left(n+\frac{23}{24}\right)}\right)$$

$$= \frac{2\pi\cdot 5^{3}}{(24\cdot 25n+24\cdot 24-1)^{\frac{1}{4}}} \cdot \sum_{\substack{k \\ (k,5)=1}} \frac{1}{5k} \frac{1}{\sqrt{5}} \sum_{\rho'=0}^{4} A_{\rho'} b_{-1,6+8\rho'} A_{k,6}^{(1)}(n) I_{\frac{1}{2}}\left(\frac{4\pi}{5k}\sqrt{\frac{1}{24}\left(25n+24-\frac{1}{24}\right)}\right)$$

$$+ \frac{2\pi\cdot 5^{3}}{(24\cdot 25n+24\cdot 24-1)^{\frac{1}{4}}} \sum_{\substack{k \\ (k,5)=1}} \frac{1}{k} A_{k,6}^{(6)}(n) I_{\frac{1}{2}}\left(\frac{4\pi}{k}\sqrt{\frac{1}{24}\left(25n+24-\frac{1}{24}\right)}\right).$$

Comparing this with (2.4) and using (4.42), we see that we shall have

$$a_n = 5^3\, p(25n + 24), \tag{5.3}$$

providing we can prove

$$A_{5k}(25n + 24) = \frac{1}{\sqrt{5}} \sum_{\rho'=0}^{4} A_{\rho'} b_{-1,6+8\rho'} A_{k,6}^{(1)}(n) \tag{5.31}$$

and

$$A_k(25n + 24) = A_{k,6}^{(6)}(n) \tag{5.32}$$

for k relatively prime to 5.

Before proving (5.31) and (5.32), we shall determine the $A_{\rho'}$ as solutions of (5.2). From (3.32) we see that the $b_{n,\rho}$ are the coefficients in the expansion

$$Q^{-\frac{1}{2}\rho} x^{\frac{1}{24}(\rho-Q\rho-Q)-\beta_\rho} \frac{\prod_{n=1}^{\infty}(1-x^n)^\rho}{\prod_{n=1}^{\infty}(1-x^{Qn})^{\rho+1}} = \sum_{n=-\mu_\rho}^{\infty} b_{n,\rho} x^n. \tag{5.4}$$

In our case, using Euler's formula, we have

$$\sum_{n=-\rho'-2}^{\infty} b_{n,6+8\rho'} x^n = 5^{-3-3\rho'} x^{-2-\rho'} \frac{(1-x-x^2+x^5+x^7-\cdots)^{6+8\rho'}}{(1-x^5-x^{10}+\cdots)^{7+8\rho'}},$$

from which we find

$$b_{-1,6} = -6\cdot 5^{-3}, \qquad b_{-2,6} = +5^{-3},$$

$$b_{-1,12} = +54\cdot 5^{-6}, \qquad b_{-2,12} = -12\cdot 5^{-6}, \qquad b_{-3,12} = +5^{-6},$$

$$b_{-1,18} = -510\cdot 5^{-9}, \quad b_{-2,18} = +135\cdot 5^{-9}, \quad b_{-3,18} = -18\cdot 5^{-9},$$
$$b_{-1,24} = +4830\cdot 5^{-12}, \quad b_{-2,24} = -1472\cdot 5^{-12}, \quad b_{-3,24} = +252\cdot 5^{-12},$$
$$b_{-1,30} = -45005\cdot 5^{-15}, \quad b_{-2,30} = +15660\cdot 5^{-15}, \quad b_{-3,30} = -3190\cdot 5^{-15},$$
$$b_{-4,18} = +5^{-9},$$
$$b_{-4,24} = -24\cdot 5^{-12}, \quad b_{-5,24} = +5^{-12},$$
$$b_{-4,30} = +405\cdot 5^{-15}, \quad b_{-5,30} = -30\cdot 5^{-15}, \quad b_{-6,30} = +5^{-15},$$

and hence

(5.51) $A_0 = 63\cdot 5^5, \quad A_1 = 52\cdot 5^8, \quad A_2 = 63\cdot 5^{10}, \quad A_3 = 6\cdot 5^{13}, \quad A_4 = 5^{15},$

and

$$\sum_{\rho'=0}^{4} A_{\rho'} b_{-1,6+6\rho'} = -5. \tag{5.52}$$

We now turn to the proof of (5.31) and (5.32). By (3.42) we now have

$$A_{k,6}^{(\nu)}(n) = \sum_{\substack{h \bmod k \\ (h,k)=1}} \frac{\omega_{h,k}^{7}}{\omega_{6h,k}^{6}} \exp\{-2\pi i k^{-1}[nh + (\nu-2)h']\}$$

and hence, by (4.2),

$$A_{k,6}^{(\nu)}(n) = \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{h,k} \exp\{-2\pi i k^{-1}[(n+1)h + (\nu-1)h']\}.$$

Taking $\nu = 6$ we see that this reduces (5.32) to

$$A_k(25n+24) = \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{h,k} \exp\{-2\pi i k^{-1}[(n+1)h + 5h']\}$$

which is merely (4.44) with $Q = 5$ and $S = 24$. Similarly, taking $\nu = 1$, we reduce (5.31) to the form

$$A_{6k}(25n+24) = \frac{1}{\sqrt{5}} \sum_{\rho'=0}^{4} A_{\rho'} b_{-1,6+6\rho'} \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{h,k} \exp[-2\pi i k^{-1}(n+1)h]$$
$$= -\sqrt{5}\, A_k(n+1),$$

where we have made use of (5.52) and (4.41). Since this is indeed a valid equation by (4.7), we have completed the proof of (5.31) and (5.32) and hence we have (5.3). Then by (5.1) we have the identity

$$e^{2\pi i \frac{11}{24}\tau} \sum_{n=0}^{\infty} p(25n+24) e^{2\pi i n\tau} = \frac{1}{5^3} \sum_{\rho'=0}^{4} A_{\rho'} F_{6+6\rho'}(\tau)$$

which, by virtue of (5.51) and (2.2) is equivalent to (1.4) and hence to (1.13).

The choice $\rho_0 = 5$ leads to the values

$$\alpha_\rho = \tfrac{1}{24}, \qquad \beta_\rho = \tfrac{23}{24}, \qquad \mu_\rho = \rho' + 2,$$

and the conditions of §2 become

$$R \equiv 5 \pmod{24}, \qquad \nu_1 = \frac{5R+23}{24}, \qquad \nu_2 = \frac{R+115}{120}.$$

This allows us to take for R any odd power of 5. For example, if $R = 5$ we have $\nu_1 = 2$ and $\nu_2 = 1$, so we need only the single function $F_5(\tau)$. With the aid of (4.2), (4.45) and (4.6) this can easily be shown to lead to Ramanujan's identity (1.11). For the value $R = 125$ we would have $\nu_1 = 27$ and $\nu_2 = 2$, so we would consider a function

$$F(\tau) = \sum_{\rho'=0}^{25} A_{\rho'} F_{5+6\rho'}(\tau)$$

and determine the $A_{\rho'}$ as solutions of

$$A_{25} b_{-27,125} = 1, \qquad \sum_{\nu-2 \le \rho' \le 25} A_{\rho'} b_{-\nu,5+6\rho'} = 0 \quad \text{for } \nu = 1, 3, 4, 5, \cdots, 26.$$

Because of the difficulties involved in the computation of the $b_{-\nu,\rho}$ and the $A_{\rho'}$ we shall not carry this case through. However, it may be mentioned that the $A_{k,\rho}^{(\nu)}(n)$ may be treated just as before. That is, by using the methods and results of §4, we can show that we have

$$A_k(125n + 99) = A_{k,5}^{(27)}(n),$$
$$A_{5k}(125n + 99) = -\sqrt{5}\, A_{k,5}^{(2)}(n).$$

Therefore the proof of an identity of the form

$$\sum_{n=0}^{\infty} p(125n + 99)x^n = \sum_{\rho'=0}^{25} B_{\rho'} x^{\rho'} \frac{\varphi(x^5)^{5+6\rho'}}{\varphi(x)^{6+6\rho'}}$$

depends only on determining the $A_{\rho'}$ and then finding

$$\sum_{\rho'=0}^{25} A_{\rho'} b_{-2,5+6\rho'} = -5.$$

6. **The case $Q = 7$.** In this and the following section we shall limit ourselves to the cases that lead to the identities (1.14) and (1.15) and shall not discuss the possibilities of other identities. For $Q = 7$ we take $\rho_0 = 4$, $\rho = 4 + 4\rho'$, and $R = 49$. We then have

$$\alpha_\rho = \tfrac{23}{24}, \qquad \beta_\rho = \tfrac{17}{24}, \qquad \mu_\rho = \rho' + 2, \qquad \nu_1 = 15, \qquad \nu_2 = 1, \qquad S = 47,$$

so we choose

$$F(\tau) = \sum_{\rho'=0}^{13} A_{\rho'} F_{4+4\rho'}(\tau) = e^{2\pi i \frac{47}{24}\tau} \sum_{n=0}^{\infty} a_n e^{2\pi i n\tau} \tag{6.11}$$

with

$$A_{13} b_{-15,56} = 1, \qquad \sum_{\nu-2 \le \rho' \le 13} A_{\rho'} b_{-\nu,4+4\rho'} = 0 \quad \text{for } \nu = 2, 3, 4, \cdots, 14.$$

By (3.11), (2.32) and (4.1) we then have

$$a_n = \frac{2\pi \cdot 7^3}{(24\cdot 49n + 24\cdot 47 - 1)^{\frac{1}{4}}} \sum_{\substack{k \\ (k,7)=1}} \frac{1}{7k}\frac{1}{\sqrt{7}}$$
$$\cdot \sum_{\rho'=0}^{13} A_{\rho'} b_{-1,4+4\rho'} A_{k,4}^{(1)}(n) I_{\frac{1}{2}}\left(\frac{4\pi}{7k}\sqrt{\frac{1}{24}\left(49n + 47 - \frac{1}{24}\right)}\right)$$
$$+ \frac{2\pi \cdot 7^3}{(24\cdot 49n + 24\cdot 47 - 1)^{\frac{1}{4}}} \sum_{\substack{k \\ (k,7)=1}} \frac{1}{k} A_{k,4}^{(15)}(n) I_{\frac{1}{2}}\left(\frac{4\pi}{k}\sqrt{\frac{1}{24}\left(49n + 47 - \frac{1}{24}\right)}\right)$$

and then, by (2.4) and (4.42),

$$a_n = 7^3\, p(49n + 47) \tag{6.12}$$

providing we can show

$$A_{7k}(49n + 47) = \frac{1}{\sqrt{7}} \sum_{\rho'=0}^{13} A_{\rho'} b_{-1,4+4\rho'} A_{k,4}^{(1)}(n), \tag{6.21}$$

$$A_k(49n + 47) = A_{k,4}^{(15)}(n). \tag{6.22}$$

The $b_{n,\rho}$ can again be determined from (5.4). It was found convenient to make use of Jacobi's formula

$$\prod_{n=1}^{\infty} (1 - x^n)^3 = 1 - 3x + 5x^3 - 7x^6 + 9x^{10} - 11x^{15} + \cdots$$

and to raise it to various powers. The desired values $b_{-r,4+4\rho'}$ could then be fairly easily found by further multiplications using Euler's formula. The $A_{\rho'}$ were then found to have the values

$$A_0 = 873278\cdot 7^2, \quad A_1 = 16784362\cdot 7^4, \quad A_2 = 69530461\cdot 7^6,$$
$$A_3 = 121951200\cdot 7^8, \quad A_4 = 117327462\cdot 7^{10}, \quad A_5 = 70415303\cdot 7^{12},$$
$$A_6 = 28303191\cdot 7^{14}, \quad A_7 = 7925456\cdot 7^{16}, \quad A_8 = 1574664\cdot 7^{18},$$
$$A_9 = 222138\cdot 7^{20}, \quad A_{10} = 21840\cdot 7^{22}, \quad A_{11} = 1428\cdot 7^{24}, \quad A_{12} = 56\cdot 7^{26}, \quad A_{13} = 7^{28},$$

while the sum in (6.21) was found to be

$$\sum_{\rho'=0}^{13} A_{\rho'} b_{-1,4+4\rho'} = -7. \tag{6.3}$$

We can now prove (6.21) and (6.22). By (3.42) and (4.2) we have

$$A_{k,4}^{(\nu)}(n) = \sum_{\substack{h \bmod k \\ (h,k)=1}} \frac{\omega_{h,k}^5}{\omega_{7h,k}^4} \exp\{-2\pi i k^{-1}[nh + (\nu - 2)h']\}$$
$$= \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{h,k} \exp\{-2\pi i k^{-1}[(n+1)h + (\nu - 1)h']\}.$$

This reduces (6.22) directly to (4.44) for $Q = 7$ and $R = 47$. Also it reduces (6.21) to

$$A_{7k}(49n + 47) = \frac{1}{\sqrt{7}} \sum_{\rho'=0}^{13} A_{\rho'} b_{-1,4+4\rho'} \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{h,k} \exp\left[-2\pi i k^{-1}(n+1)h\right]$$

which, by (4.41) and (6.3), is merely (4.7) for the same values of Q and R. This completes the proof of (6.12) which can easily be shown, with the aid of (6.11) and the values of $A_{\rho'}$, to yield the identity (1.14).

7. **The case $Q = 13$.** We now take $Q = 13$, $\rho_0 = 1$, $\rho = 1 + 2\rho'$, $R = 13$ and have

$$\alpha_\rho = \tfrac{11}{24}, \qquad \beta_\rho = \tfrac{23}{24}, \qquad \mu_\rho = \rho' + 2, \qquad \nu_1 = 8, \qquad \nu_2 = 1, \qquad S = 6.$$

Hence we take

$$F(\tau) = \sum_{\rho'=0}^{6} A_{\rho'} F_{1+2\rho'}(\tau) = e^{2\pi i \frac{11}{24}\tau} \sum_{n=0}^{\infty} a_n e^{2\pi i n \tau} \tag{7.11}$$

with

$$A_6 b_{-8,13} = 1, \qquad \sum_{\nu-2 \le \rho' \le 6} A_{\rho'} b_{-\nu,1+2\rho'} = 0 \quad \text{for } \nu = 2, 3, 4, \cdots, 7,$$

and have as before

$$a_n = \frac{2\pi \cdot 13^{\frac{1}{4}}}{(24 \cdot 13n + 24 \cdot 6 - 1)^{\frac{3}{4}}} \sum_{\substack{k \\ (k,13)=1}} \frac{1}{13k} \frac{1}{\sqrt{13}}$$
$$\cdot \sum_{\rho'=0}^{6} A_{\rho'} b_{-1,1+2\rho'} A_{k,1}^{(1)}(n) I_{\frac{3}{2}}\left(\frac{4\pi}{13k}\sqrt{\frac{1}{24}\left(13n + 6 - \frac{1}{24}\right)}\right)$$
$$+ \frac{2\pi \cdot 13^{\frac{1}{4}}}{(24 \cdot 13n + 24 \cdot 6 - 1)^{\frac{3}{4}}} \sum_{\substack{k \\ (k,13)=1}} \frac{1}{k} A_{k,1}^{(8)}(n) I_{\frac{3}{2}}\left(\frac{4\pi}{k}\sqrt{\frac{1}{24}\left(13n + 6 - \frac{1}{24}\right)}\right),$$

and therefore

$$a_n = 13^{\frac{1}{4}} p(13n + 6), \tag{7.12}$$

providing we have

$$A_{13k}(13n + 6) = \frac{1}{\sqrt{13}} \sum_{\rho'=0}^{6} A_{\rho'} b_{-1,1+2\rho'} A_{k,1}^{(1)}(n), \tag{7.21}$$

$$A_k(13n + 6) = A_{k,1}^{(8)}(n). \tag{7.22}$$

The $b_{n,\rho}$ can easily be found from (5.4) and the $A_{\rho'}$ are then found to be

$$A_0 = 143 \cdot 13^{\frac{1}{4}}, \qquad A_1 = 468 \cdot 13^{\frac{1}{4}}, \qquad A_2 = 494 \cdot 13^{\frac{1}{4}}, \qquad A_3 = 260 \cdot 13^{\frac{1}{4}},$$
$$A_4 = 78 \cdot 13^{\frac{3}{4}}, \qquad A_5 = 13 \cdot 13^{\frac{3}{4}}, \qquad A_6 = 13^{\frac{3}{4}},$$

while the sum in (7.21) is found to have the value

$$\sum_{\rho'=0}^{6} A_{\rho'} b_{-1,1+2\rho'} = 13. \tag{7.3}$$

With these values, (7.12) and (7.11) will then give us the identity (1.15).

There remains only to prove (7.21) and (7.22). Again by (3.42) and (4.2) we have

$$\begin{aligned} A_{k,1}^{(\nu)}(n) &= \sum_{\substack{h \bmod k \\ (h,k)=1}} \frac{\omega_{h,k}^2}{\omega_{13h,k}} \exp\{-2\pi i k^{-1}[nh + (\nu-2)h']\} \\ &= \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{13h,k} \exp\{-2\pi i k^{-1}[(n+1)h + (\nu-1)h']\}. \end{aligned} \tag{7.4}$$

Then (7.22) is equivalent to

$$A_k(13n+6) = \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{13h,k} \exp\{-2\pi i k^{-1}[(n+1)h + 7h']\}$$

which is merely (4.45) with $Q = 13$, $S = 6$. Also (7.3) and (7.4) reduce (7.21) to

$$A_{13k}(13n+6) = \sqrt{13} \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{13h,k} \exp[-2\pi i k^{-1}(n+1)h].$$

If we now let $\bar{k}$ again be any solution of $k\bar{k} \equiv 1 \pmod{13}$, we can write the above in the form

$$\begin{aligned} A_{13k}(13n+6) &= \sqrt{13} \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{13h,k} \exp[-2\pi i k^{-1}(n+1)(1-k\bar{k})h] \\ &= \sqrt{13} \sum_{\substack{h \bmod k \\ (h,k)=1}} \omega_{h,k} \exp\left[-2\pi i k^{-1}(n+1)\frac{1-k\bar{k}}{13}h\right] \\ &= \sqrt{13}\, A_k\left[(n+1)\frac{1-k\bar{k}}{13}\right] \end{aligned}$$

which is a true equation by (4.6). This then completes the proof of (1.15).

8. **The divisibility properties.** The methods employed by Ramanujan to obtain divisibility properties from his identities can now be applied to (1.13) and (1.14) to find new ones. The identity (1.15) does not appear to be able to yield any such simple properties. We shall make use of Euler's and Jacobi's formulas

$$\prod_{n=1}^{\infty}(1-x^n) = \sum_{n=-\infty}^{\infty}(-1)^n x^{\frac{1}{2}n(3n+1)} \tag{8.11}$$

and

$$\prod_{n=1}^{\infty}(1-x^n)^3 = \sum_{n=0}^{\infty}(-1)^n(2n+1)x^{\frac{1}{2}n(n+1)} \tag{8.12}$$

and the congruence property

$$\prod_{n=1}^{\infty}(1-x^{Qn}) \equiv \prod_{n=1}^{\infty}(1-x^n)^Q \pmod{Q}. \tag{8.13}$$

This last property follows immediately from the binomial expansion of $(1-x^n)^Q$ since we are supposing that Q is a prime.

By (1.13) we have

$$\frac{1}{25}\sum_{n=0}^{\infty}p(25n+24)x^n \equiv 63\,\frac{\varphi(x^5)^6}{\varphi(x)^7} \pmod 5$$

and hence, by (8.13),

$$\frac{1}{25}\sum_{n=0}^{\infty}p(25n+24)x^n \equiv 3\varphi(x^5)^3\varphi(x^5)\varphi(x)^3 \pmod 5.$$

Now using (8.11) and (8.12) we have

$$\frac{1}{25}\sum_{n=0}^{\infty}p(25n+24)x^n \equiv 3\sum_{j=0}^{\infty}(-1)^j(2j+1)x^{\frac{1}{2}j(j+1)} \cdot \sum_{k=-\infty}^{\infty}(-1)^k x^{\frac{1}{2}k(3k+1)}\sum_{l=0}^{\infty}(-1)^l(2l+1)x^{\frac{1}{2}l(l+1)} \pmod 5$$

and therefore

$$\tfrac{1}{25}p(25n+24) \equiv 3\sum(-1)^{j+k+l}(2j+1)(2l+1) \pmod 5, \tag{8.2}$$

where the sum extends over all values j, k, l with $j \geqq 0, l \geqq 0$ and

$$\tfrac{5}{2}j(j+1)+\tfrac{5}{2}k(3k+1)+\tfrac{1}{2}l(l+1) = n$$

which may be written as

$$15(2j+1)^2+5(6k+1)^2+3(2l+1)^2 = 24n+23.$$

Therefore in the sum (8.2) all the terms will be multiples of 5 if $24n+23$ is a multiple of 5 or there will be no terms at all if $24n+23$ is congruent to 3 times a non-residue modulo 5, that is,

$$24n+23 \equiv 0, 1, 4 \pmod 5,$$

$$n \equiv 3, 2, 4 \pmod 5.$$

For these values of n we have

$$25n+24 \equiv 99, 74, 124 \pmod{125}$$

and, by (8.2),

$$p(25n+24) \equiv 0 \pmod{125},$$

so we have proved (1.23).

In a similar way we prove (1.25). By (1.14), (8.11), (8.12), and (8.13) we have

$$\frac{1}{49}\sum_{n=0}^{\infty} p(49n+47)x^n \equiv 2546\,\frac{\varphi(x^7)^4}{\varphi(x)^6} \equiv 5\varphi(x^7)^3\varphi(x)^2 \pmod 7$$

$$\equiv 5\sum_{j=0}^{\infty}(-1)^j(2j+1)x^{\frac{1}{2}j(j+1)}\sum_{k=-\infty}^{\infty}(-1)^k x^{\frac{1}{2}k(3k+1)}\sum_{l=-\infty}^{\infty}(-1)^l x^{\frac{1}{2}l(3l+1)} \pmod 7,$$

and therefore

$$\frac{1}{49}p(49n+47) \equiv 5\sum(-1)^{j+k+l}(2j+1) \pmod 7, \tag{8.31}$$

where the sum extends over all $j \geq 0$ and all k and l satisfying the equation

$$\tfrac{1}{2}j(j+1) + \tfrac{1}{2}k(3k+1) + \tfrac{1}{2}l(3l+1) = n$$

which may be written as

$$21(2j+1)^2 + (6k+1)^2 + (6l+1)^2 = 24n + 23. \tag{8.32}$$

Now if $24n + 23 \equiv 0 \pmod 7$, we have

$$(6k+1)^2 + (6l+1)^2 \equiv 0 \pmod 7,$$

and hence

$$6k + 1 \equiv 6l + 1 \equiv 0 \pmod 7$$

since the quadratic residues modulo 7 are 0, 1, 2, and 4. Then from (8.32) we have

$$3(2j+1)^2 \equiv \frac{24n+23}{7} \pmod 7,$$

and therefore, by (8.31),

$$p(49n+47) \equiv 0 \pmod{343} \tag{8.33}$$

for all n such that

$$\frac{24n+23}{7} \equiv 0, 1, 2, 4 \pmod 7;$$

that is,

$$n \equiv 46, 32, 18, 39 \pmod{49}. \tag{8.34}$$

This completes the proof of (1.25), which is equivalent to (8.33) and (8.34).

The proof of (1.24) is somewhat different. We first discuss the coefficients b_n in the expansion

$$\varphi(x)^{18} = \sum_{n=0}^{\infty} b_n x^n. \tag{8.41}$$

By (1.3) we have

$$\eta(\tau)^{18} = e^{\frac{1}{2}\pi i\tau}\varphi(e^{2\pi i\tau})^{18} = e^{\frac{1}{2}\pi i\tau}\sum_{n=0}^{\infty} b_n e^{2\pi i n\tau}, \tag{8.42}$$

and hence

$$\begin{aligned}\sum_{\rho=0}^{4}\eta\left(\frac{\tau+24\rho}{5}\right)^{18} &= \sum_{\rho=0}^{4} e^{\frac{18}{5}\pi i(\tau+24\rho)}\sum_{n=0}^{\infty} b_n e^{\frac{1}{2}\pi i n(\tau+24\rho)}\\ &= e^{\frac{18}{5}\pi i\tau}\sum_{n=0}^{\infty} b_n e^{\frac{1}{2}\pi i n\tau}\sum_{\rho=0}^{4} e^{\frac{1}{2}\pi i(24n+18)\rho}\\ &= 5e^{\frac{18}{5}\pi i\tau}\sum_{n=0}^{\infty} b_{5n+3}e^{\frac{1}{2}\pi i(5n+3)\tau}\\ &= 5e^{\frac{1}{2}\pi i\tau}\sum_{n=0}^{\infty} b_{5n+3}e^{2\pi i n\tau}.\end{aligned} \tag{8.43}$$

Now the function

$$H(\tau) = 5^9\eta(5\tau)^{18} + \sum_{\rho=0}^{4}\eta\left(\frac{\tau+24\rho}{5}\right)^{18} \tag{8.44}$$

is a modular form belonging to the full modular group. To show this we need only exhibit its behavior under the two generators of the modular group

$$\tau' = \tau + 1, \qquad \tau' = \frac{-1}{\tau}.$$

By (8.42) we have

$$\eta(\tau+5)^{18} = e^{\frac{1}{2}\pi i}\eta(\tau)^{18},$$

and therefore we have

$$\begin{aligned}H(\tau+1) &= 5^9\eta(5\tau+5)^{18} + \sum_{\rho=0}^{4}\eta\left(\frac{\tau+1+24\rho}{5}\right)^{18}\\ &= 5^9\eta(5\tau+5)^{18} + \sum_{\rho=0}^{4}\eta\left(\frac{\tau+24(\rho-1)}{5}+5\right)^{18}\\ &= 5^9 e^{\frac{1}{2}\pi i}\eta(5\tau)^{18} + e^{\frac{1}{2}\pi i}\sum_{\rho=0}^{4}\eta\left(\frac{\tau+24(\rho-1)}{5}\right)^{18}\\ &= 5^9 e^{\frac{1}{2}\pi i}\eta(5\tau)^{18} + e^{\frac{1}{2}\pi i}\sum_{\rho=1}^{4}\eta\left(\frac{\tau+24(\rho-1)}{5}\right)^{18} + e^{\frac{1}{2}\pi i}\eta\left(\frac{\tau+24\cdot 4}{5}-24\right)^{18}\\ &= e^{\frac{1}{2}\pi i}H(\tau)\end{aligned}$$

since $\eta(\tau-24)^{18} = \eta(\tau)^{18}$ by (8.42). For the second transformation we have

$$\begin{aligned}H\left(\frac{-1}{\tau}\right) &= 5^9\eta\left(\frac{-5}{\tau}\right)^{18} + \sum_{\rho=0}^{4}\eta\left(\frac{-\frac{1}{\tau}+24\rho}{5}\right)^{18}\\ &= 5^9\eta\left(\frac{-1}{\frac{1}{5}\tau}\right)^{18} + \eta\left(\frac{-1}{5\tau}\right)^{18} + \sum_{\rho=1}^{4}\eta\left(\frac{24\rho\,\frac{\tau+24\sigma}{5} - \frac{1+24^2\rho\sigma}{5}}{5\,\frac{\tau+24\sigma}{5} - 24\sigma}\right)^{18},\end{aligned}$$

where we shall define σ by $24^2\rho\sigma \equiv -1 \pmod 5$, $0 < \sigma < 5$. We now use (3.2) and (4.32) to find

$$H\left(\frac{-1}{\tau}\right) = (-i\tau)^9\left\{\eta(\tfrac{1}{5}\tau)^{18} + 5^9\eta(5\tau)^{18} + \sum_{\sigma=1}^{4}\eta\left(\frac{\tau+24\sigma}{5}\right)^{18}\right\}$$
$$= (-i\tau)^9 H(\tau).$$

Also we have from (3.2) and (4.32)

$$\eta(\tau+1)^{18} = e^{\frac{3}{4}\pi i}\eta(\tau)^{18}, \qquad \eta\left(\frac{-1}{\tau}\right)^{18} = (-i\tau)^9\eta(\tau)^{18}$$

and hence the function $H(\tau)/\eta(\tau)^{18}$ is a modular function. The only part of the fundamental region at which $H(\tau)/\eta(\tau)^{18}$ might not be bounded is the point $\tau = i\infty$, but there we have by (8.42), (8.43) and (8.44)

$$\frac{H(\tau)}{\eta(\tau)^{18}} = \frac{5^9 e^{\frac{1}{2}\pi i\tau}\sum_{n=0}^{\infty} b_n e^{10\pi i n\tau} + 5e^{\frac{1}{2}\pi i\tau}\sum_{n=0}^{\infty} b_{5n+3}e^{2\pi i n\tau}}{e^{\frac{1}{2}\pi i\tau}\sum_{n=0}^{\infty} b_n e^{2\pi i n\tau}}$$
$$= \frac{5b_3}{b_0} + \frac{5}{b_0^2}(b_0b_8 - b_1)e^{2\pi i\tau} + \cdots.$$

Therefore $H(\tau)/\eta(\tau)^{18}$ is bounded at $i\infty$ and approaches the value $5b_3/b_0 = -2550$. Since then the function is a bounded modular function, it is a constant and its value is -2550. We therefore have

$$H(\tau) = -2550\,\eta(\tau)^{18},$$

and then, by (8.42), (8.43) and (8.44),

$$5^8x^3\sum_{n=0}^{\infty} b_n x^{5n} + \sum_{n=0}^{\infty} b_{5n+3}x^n = -510\sum_{n=0}^{\infty} b_n x^n.$$

Equating coefficients of like powers of x gives us

$$b_{5n+3} \equiv -510\,b_n \pmod{5^8} \tag{8.5}$$

which we shall use below.

From (8.13) we have

$$\varphi(x^5)^5 \equiv \varphi(x)^{25} \pmod{25},$$

and therefore, from (1.13),

$$\frac{1}{25}\sum_{n=0}^{\infty} p(25n+24)x^n \equiv 63\,\frac{\varphi(x^5)^6}{\varphi(x)^7} \pmod{25}$$
$$\equiv 13\,\varphi(x)^{18}\varphi(x^5) \pmod{25}.$$

Using the notation

$$\varphi(x) = \sum_{n=0}^{\infty} a_n x^n$$

and (8.41), we have

$$\frac{1}{25}\sum_{n=0}^{\infty} p(25n+24)x^n \equiv 13\sum_{j=0}^{\infty} b_j x^j \sum_{k=0}^{\infty} a_k x^{5k} \pmod{25},$$

and hence

$$\tfrac{1}{25}p(25n+24) \equiv 13\sum_{0\le k\le \frac{1}{5}n} a_k b_{n-5k} \pmod{25}.$$

We replace n by $5n+3$ and use (8.5) to obtain

$$\frac{1}{25}p(125n+99) \equiv 13\sum_{k=0}^{n} a_k b_{5n-5k+3} \pmod{25}$$

$$\equiv -13\cdot 510\sum_{k=0}^{n} a_k b_{n-k} \pmod{25}.$$

Therefore we have

$$\frac{1}{125}\sum_{n=0}^{\infty} p(125n+99)x^n \equiv 4\sum_{n=0}^{\infty}\sum_{k=0}^{n} a_k b_{n-k}x^n \pmod{5}$$

$$\equiv 4\sum_{k=0}^{\infty} a_k x^k \sum_{j=0}^{\infty} b_j x^j \pmod{5}$$

$$\equiv 4\,\varphi(x)^{19} \pmod{5}.$$

Now by (8.13) and (1.11) we have

$$\varphi(x)^{19} \equiv \frac{\varphi(x^5)^5}{\varphi(x)^5} \equiv \frac{1}{5}\sum_{n=0}^{\infty} p(5n+4)\,x^n \pmod{5}$$

and hence

$$\frac{1}{125}p(125n+99) \equiv \frac{4}{5}p(5n+4) \pmod{5}.$$

Replacing n by $5n+4$, we find

$$\frac{1}{125}p(625n+599) \equiv \frac{4}{5}p(25n+24) \pmod{5},$$

and therefore (1.24) since $p(25n+24) \equiv 0 \pmod{25}$.

UNIVERSITY OF PENNSYLVANIA.

HEREDITARY ARC SUMS

By O. G. Harrold, Jr.

1. The object of this paper is to obtain conditions under which a compact locally connected continuum and all of its subcontinua are arc sums. A continuous curve possessing this property will be called a *hereditary arc sum*. G. T. Whyburn[1] has shown that a necessary and sufficient condition that a curve be an arc sum, that is, the sum of a countable number of simple continuous arcs, is that it have a countable number of end-points and that each true cyclic element be an arc sum. It is well known that if an acyclic curve is an arc sum, every curve contained in it is an arc sum. The corresponding proposition about continuous curves in general, or even for the class of cyclic curves, is not true, as is easily shown by examples.

It is proved in this paper that the property of being a hereditary arc sum is equivalent to each of the following:

(1) *each acyclic curve in the given curve is an arc sum*;
(2) *the boundary of each region in the curve is a countable point set*;
(3) *every connected subset is a* G_δ;
(4) *the non-local separating points are countable*;
(5) *for every connected subset* G, $\bar{G} - G$ *is countable.*

It is shown that every curve containing a convergence continuum contains an acyclic non-arc-sum, hence that the class of hereditary arc sums is contained in the class of hereditarily locally connected continua. In §5 it is proved that the property of being a hereditary arc sum is invariant under continuous transformations which are either monotone or interior.

The author wishes to express his appreciation to Professor G. T. Whyburn, who has very generously made most helpful suggestions during the preparation of this paper.

2. It is evident that each acyclic curve in M is an arc sum if M is a hereditary arc sum. To establish the converse we prove first

Theorem A. *Every hereditarily locally connected compact continuum which is not an arc sum contains an acyclic non-arc-sum.*

Proof. It will be convenient to introduce the notion of an *adjacent cycle point*. A point p is called an adjacent cycle point of the set M provided that every neighborhood of p contains a simple closed curve lying in M. Clearly, the set of adjacent cycle points is closed. Let M_1 be the set of adjacent cycle

Received September 12, 1938.

[1] G. T. Whyburn, *Concerning continuous curves and arc sums*, Fundamenta Mathematicae, vol. 14(1929), pp. 103–106.

points of M, and in general, if the ordinal β has a predecessor, let M_β be the set of adjacent cycle points of $M_{\beta-1}$. If the ordinal β has no immediate predecessor, set $M_\beta = \prod_{i<\beta} M_i$. Each M_α is closed, and, if $\alpha > \beta$, then $M_\alpha \subset M_\beta$. If some M_β is compact, then $M_\beta = M_{\beta+1}$ for all ordinals $\beta > \beta_0$.

Let M now denote the given curve which by assumption is not an arc sum. For each ordinal i of the first or second class set $M = M_i + U_i$, $U_i = M - M_i$. Suppose that for some ordinal $M_\beta = 0$. If we let β denote the smallest ordinal for which this is true, it is clear that β has a predecessor $\beta - 1 = \alpha$. Setting $M = M_\alpha + U_\alpha$ ($M_{\alpha+1} = 0$), we shall prove that M_α has only a finite number of distinct topological circles. Since M is hereditarily locally connected, M_α has only a finite number of components of diameter $\geqq \epsilon > 0$. Suppose now that $C_1, C_2, C_3, \cdots$ is a sequence of distinct simple closed curves in M_α. If there is a subsequence with diameter not tending to zero, we get a contradiction, for then there must be infinitely many C_k in some one component W of M_α. But W is a continuous curve in M_α which clearly is no baum im kleinen and hence contains an adjacent cycle point. If there is a subsequence of the C_i with diameter tending to zero, we get a contradiction. For, let $c_k \in C_k$; then if p is a limit point of $\sum c_k$, $p \in M_{\alpha+1}$, but $M_{\alpha+1} = 0$. Thus if the topological circles in M are $C_1, C_2, \cdots, C_n$, $M_\alpha - (C_1 + C_2 + \cdots + C_n)$ is the sum of a countable number of closed sets

$$T_k = M_\alpha - V_{1/k}(C_1 + C_2 + \cdots + C_n) \quad (k = 1, 2, 3, \cdots).$$

Each T_k has only point and acyclic curve components. Only a finite number have diameter greater than a given positive number. Hence each T_k is contained in an acyclic curve S_k in M.[2] Thus $M_\alpha \subset (C_1 + C_2 + \cdots + C_n) + \sum S_k$. If some S_k is a non-arc-sum, our end is attained; if not, M_α is contained in an arc sum in M.

The set $U_\alpha = M - M_\alpha$ may be covered by an acyclic curve sum as follows. If $M = M_1$, then $M_1 = M_2$, and no $M_\beta = 0$. Thus $U_1 \neq 0$. If $p \in U_1$, then p belongs to a neighborhood V such that $M\bar{V}$ has only point and acyclic curve components with only a finite number of the latter having a diameter greater than a given positive number, so that $M\bar{V}$ is contained in an acyclic curve in M. By the Lindelöf Theorem U_1 is covered by a countable collection of such acyclic curves. If one of these is a non-arc-sum, our end is attained; if not, U_1 is covered by an arc sum. If U_1 be covered by an arc sum, then $U_{i+1} = U_i + (M_i - M_{i+1})$ is covered by an arc sum, or our end is attained; for $M_i - M_{i+1}$ can be covered by an acyclic curve sum as before. Thus, if $U_1, U_2, \cdots$ are all covered by arc sums, so is $U = \sum U_i$. Proceeding in this way, we arrive at a first ordinal β such that U_β is no longer coverable by an arc sum in M, or all U_i are contained in arc sums in M. In the first case, since β evidently has a predecessor, $M_{\beta-1} - M_\beta$ is covered by an acyclic curve sum one of whose

[2] L. Zippin, *On continuous curves irreducible about subsets*, Fundamenta Mathematicae, vol. 20(1933), pp. 197–205.

elements is an acyclic non-arc-sum. In the second case, since $M = M_\alpha + U_\alpha$, we have a contradiction to the fact that M is no arc sum.

The desired type of curve has been shown to exist in M provided $M_{\alpha+1} = 0$ for some ordinal of the first or second class. If this is not the case, $M_\alpha = M_{\alpha+1} \neq 0$ for all ordinals $\alpha > \alpha_0$. The truth of the proposition in this circumstance follows from

LEMMA B. *Every compact locally connected continuum which contains a subset N such that $N = N_1$ contains an acyclic non-arc-sum.*

Proof. We shall prove this lemma at present only for hereditarily locally connected continua. This is sufficient for the completion of Theorem A. That the conclusion holds if M contains continua of convergence will be a consequence of the Theorem D, which is independent of Theorem A and Lemma B.

Let b_0 be any point of N and $U(b_0)$ an open connected subset of M containing b_0. Let J be one of the infinitely many topological circles of N in U. Let t_1^1 be any arc in J with end-points a_1 and a_2. Define $C_1 = t_1^1$. Let W_i be an open connected set in M containing a_i with the following properties:

(i) the diameter of W_i is less than a third of the distance between a_1 and a_2, $\delta(W_i) < \frac{1}{3}\rho(a_1, a_2)$;

(ii) $B(W_i)C_1$ consists of a single point.

Let b_i be the unique point in W_iC_1 such that it is the first point along C_1 for which $\rho = \epsilon$ as C_1 is traversed from $B(W_i)C_1$ to a_i, where $3\epsilon > 0$, $3\epsilon < \min \rho(a_i, B(W_i))$. Let $U_1(b_1)$ and $U_2(b_2)$ be open connected sets in M, $\delta(U_i) < \epsilon$. In U_i let x_i be an arc in N for which $x_iC_1 = 0$. This is possible since $b_i \in N$. Let y_i be an arc in U_i joining x_i and C_1. The continuous curve $x_i + y_i$ contains an arc t_i^2 containing a subarc of N in x_i, and one end-point of t_i^2 is in this subarc and one in C_1. Define $C_2 = C_1 + t_1^2 + t_2^2$. The set C_2 is obviously a continuous curve which has the following properties:

(α) it is acyclic;

(β) each end-point of C_1 is an end-point of C_2;

(γ) a_{12} and a_{22}, the end-points of t_1^2 and t_2^2, respectively, not in C_1, are also end-points of C_2;

(δ) each end-point a_{ij} of C_2 is the end-point of an arc in NC_2 of diameter $\delta_{ij} > 0$;

(ϵ) to each a_{ij} of C_2 can be associated an open connected set W_{ij} such that $C_2B(W_{ij})$ consists of a single point and $\delta(W_{ij}) < \mu$, where

$$0 < 3\mu < \min [\rho(a_{ij}, \sum B(U_k) + \sum B(W_k)), \delta_{ij}, \rho(a_{i2}, C_1)]$$

for i, j running independently over 1, 2.

If we continue this process, it is easy to establish the existence of $C_1, C_2, \cdots$, where C_n is such that

(α) it is acyclic (and $C_{n-1} \subset C_n$),

(β) each end-point of C_{n-1} is an end-point of C_n,

(γ) $a_{\alpha_1 \cdots \alpha_{n-1}2}$ ($\alpha_i = 1, 2$), which are the end-points of the arcs added to C_{n-1} to get C_n which do not lie in C_{n-1}, are also end-points of C_n,

(δ) each end-point $a_{\alpha_1 \cdots \alpha_n}$ of C_n is an end-point of an arc in $C_n N$ of diameter $\delta_{\alpha_1 \cdots \alpha_n}$,

(ϵ) to each end-point $a_{\alpha_1 \cdots \alpha_n}$ of C_n is associated an open connected set $W_{\alpha_1 \cdots \alpha_n}$ such that

(a) $a_{\alpha_1 \cdots \alpha_n} \in W_{\alpha_1 \cdots \alpha_n}$,

(b) $B(W_{\alpha_1 \cdots \alpha_n}) C_n$ consists of a single point,

(c) $\delta(W_{\alpha_1 \cdots \alpha_n}) < \mu$, where 3μ is a positive number less than

$$\min [\rho(a_{\alpha_1 \cdots \alpha_n}, \sum B(W_{\alpha_1 \cdots \alpha_{n-1}}) + \sum B(U_{\alpha_1 \cdots \alpha_{n-1}})), \delta_{\alpha_1 \cdots \alpha_n}, \rho(a_{\alpha_1 \cdots \alpha_{n-1}}, C_{n-1})].$$

There is thus determined a monotone increasing sequence of acyclic curves in M: $C_1, C_2, \cdots$. Set $C = \sum C_i$. The set C is a connected subset of M; hence $\bar{C}$ is a compact continuum. If M is hereditarily locally connected, $\bar{C}$ is a continuous curve, which by the method of construction of the C_i possesses the following properties:

(1) each end-point of C_n is an end-point of $\bar{C}$;

(2) each point of $\bar{C} - C$ is an end-point of $\bar{C}$;

(3) $\bar{C}$ is acyclic;

(4) $\bar{C}$ has an uncountable number of end-points.

From the dyadic construction involved it is clear that property (4) is a consequence of (1), (2) and (3). The first three properties follow from the fact that $B(W_{\alpha_1 \cdots \alpha_n})\bar{C}$ is a single point for all n and all combinations of subscripts. Since $\bar{C}$ has an uncountable number of end-points, it is an acyclic non-arc-sum. Hence the Lemma B with the additional hypothesis that M be hereditarily locally connected has been proved. This completes the proof of Theorem A.

3. In this section we prove Theorems C, D and E which will complete the first two characterizations.

THEOREM C. *Every locally connected continuum M containing a region R whose boundary is an uncountable point set contains an acyclic non-arc-sum.*

Proof. The set $B(R)$ is an uncountable G_δ. It therefore contains a compact, perfect, totally disconnected subset B. The set $M - B$ is open and contains a component Q which contains R. Thus Q is an open connected set in M whose boundary is precisely B. Since Q has property S,[3] the points of B are regularly accessible[4] from Q.

Let a_1 and a_2 be any two distinct points in B. An arc t^0 can be found joining these points and such that $t^0 \subset Q + a_1 + a_2$. If $\rho(a_1, a_2) = d_1$, let W^1 and W^2 be neighborhoods of a_1 and a_2, respectively, such that $\delta(W^i) < \frac{1}{3} d_1$ $(i = 1, 2)$ and $B(W^i)t^0$ is a single point. Set $a_i = a_{ii}$. In $W^i - a_{ii}$ select a point a_{ij} $(i \neq j)$ in B. Since B is perfect, this is possible. Since each point of B is

[3] G. T. Whyburn, *Concerning S regions in locally connected continua*, Fundamenta Mathematicae, vol. 20(1933), pp. 131–139.

[4] B. Aitchison, *Concerning regular accessibility*, Fundamenta Mathematicae, vol. 20 (1933), pp. 117–125, Theorem 7.

regularly accessible from Q, an arc t^i can be found joining a_{ij} to t^0 which will have only a_{ij} in $M - Q$ and such that $t^i \subset W^i$. In fact, there is no loss in assuming that $t^0 \cdot t^i$ consists of a single point. Set $C^2 = t^0 + t^1 + t^2$. Let

$$d_2 = \min [\rho(a_{ij}, t^0), \rho(a_{ij}, B(W^i)), \rho(a_{ii}, a_{ij} + B(W^i))] \quad (i \neq j).$$

To each a_{ij} $(i, j = 1, 2)$ is associated a neighborhood W^{ij} of diameter $< \frac{1}{2}d_2$ such that $C^2B(W^{ij})$ is a single point. Continuing this process, we determine a sequence of curves $C^1, C^2, C^3, \cdots$ such that

(a) $C^n \subset C^{n-1} + \sum W^{\alpha_1 \cdots \alpha_{n-1}}$;

(b) all end-points of C^n are in B;

(c) to each $a_{\alpha_1 \cdots \alpha_n}$, which is an end-point of C^m $(m \geqq n)$, is associated a neighborhood $W^{\alpha_1 \cdots \alpha_n}$ such that $\delta(W^{\alpha_1 \cdots \alpha_n}) < \frac{1}{2}d_n$, where

$$d_n = \min [\rho(a_{\alpha_1 \cdots \alpha_{n-2} ii}, C^{n-1} + \sum B(W^{\alpha_1 \cdots \alpha_{n-1}})),$$

$$\rho(a_{\alpha_1 \cdots \alpha_{n-2} ii}, B(W^{\alpha_1 \cdots \alpha_{n-1}}) + a_{\alpha_1 \cdots \alpha_{n-2} ij})] \quad (\alpha_1, \alpha_2, \cdots, \alpha_n = 1, 2; i \neq j),$$

and such that $C^nB(W^{\alpha_1 \cdots \alpha_n})$ is a single point.

Define $C = \sum C^i$. Since $C^{i+1} \supset C^i$, $\bar{C}$ is a compact continuum. By the method of construction, for fixed n and all m, $C^m \subset C^n + \sum W^{\alpha_1 \cdots \alpha_n}$, and $\bar{C} \subset C^n + \sum W^{\alpha_1 \cdots \alpha_n}$. As the diameter of $W^{\alpha_1 \cdots \alpha_n}$ tends to zero as n increases indefinitely, each point x of $\bar{C} - \bar{C}B$ is contained in a neighborhood V such that $V\bar{C} = VC^n$, $n \geqq n^0$, so that $\bar{C}$ is locally connected at such points. Since B is totally disconnected, $\bar{C}$ is locally connected at each of its points and hence is a continuous curve. The set $B\bar{C}$ is a closed, totally disconnected subset of a continuous curve $\bar{C}$. Hence there exists an acyclic curve A in $\bar{C}$ which contains $\bar{C}B$. It remains only to show that A has uncountably many end-points. Each set $\Pi \overline{W}^{\alpha_1 \cdots \alpha_n}$ for a definite sequence (α_i) of indices determines a unique point in $B\bar{C}$ which is an end-point in $\bar{C}$, and so likewise in A. The number of end-points in A is clearly uncountable and therefore A is no arc sum.

COROLLARY. *If every subcontinuum of a locally connected continuum M is an arc sum, the boundary of every region in M is countable.*

THEOREM D. *Every locally connected continuum M containing a convergence continuum contains an acyclic non-arc-sum.*

Proof. Let K be a convergence continuum in M. Let $p \in K$. Let L be a subcontinuum of K such that $pL = 0$. Since M is locally connected, a region U exists in M containing p such that $\bar{U}L = 0$. The component R of $M - L$ containing U will contain continua K_i such that $L \subset \lim (K_i)$. Hence $L \subset B(R)$. Thus R has an uncountable boundary; hence applying the Theorem C to R, we see that M contains an acyclic non-arc-sum.

THEOREM E. *If the boundary of every region in the compact continuum M is countable, M is a hereditary arc sum.*

Proof. The given conditions imply that M is a continuous curve. For if (K_i) were a sequence of mutually exclusive subcontinua of M converging to the non-degenerate continuum K of M, we could find a countable subset X^0 in K

such that $X^0 + (M - K)$ is connected.[5] Since X^0 is an F_σ, $Y^0 = K - X^0$ is an uncountable G_δ. Let D be a compact, perfect subset of Y^0. Then $M - D$ is open and also connected, since D is a subset of $K - X^0$.[5] Hence $M - D$ is a region in M with an uncountable boundary, contrary to assumption.

If M were not a hereditary arc sum, it would contain a non-arc-sum, which by the Theorems A and D can be assumed to be an acyclic curve N. Let the uncountable set of end-points in N be E. Since E is an uncountable G_δ, it contains a compact, perfect subset D. The set $M - D$ is open. Let the component of this set containing $N - E$ be J. Then $D = B(J)$. But this set is uncountable. Hence M can contain no continuum which is not an arc sum. Thus we have completed the first two characterizations of hereditary arc sums mentioned in §1.

4. We shall now show that this class of continua is identical with the class of curves (M) defined by any one of the following properties:

(1) every connected subset is a G_δ;

(2) the non-local separating points of M are countable;

(3) for every connected subset G of M the set $\bar{G} - G$ is countable.

G. T. Whyburn[6] has shown that for arbitrary continua properties (1)-(3) are equivalent.

To show that the curve H with properties (1)-(3) is a hereditary arc sum merely choose G to be an arbitrary region in H. By (3) it has a countable boundary, hence by Theorem E the continuum H is a hereditary arc sum. Now, if H is a hereditary arc sum, it will suffice to show that the non-local separating points are countable. Suppose the set N of non-local separating points is uncountable. Since N is a G_δ, it contains a compact, perfect subset N^0. If $H - N^0$ is connected, we have already a contradiction since this is a region with an uncountable boundary. On the other hand, if N^0 separates x and y in H, then H, being a continuous curve, contains an irreducible cutting $N^1 \subset N^0$ between x and y. Now it is well known that N^1 is the boundary of the components X and Y of $H - N^1$ containing x and y, respectively, and thus, since H is a hereditary arc sum, N^1 must be countable. The set N^1 is countable and closed hence it contains an isolated point p. Since p is an isolated point of an irreducible cutting, it must be a local separating point in H. But $p \in N^1 \subset N^0 \subset N$, hence the assumption that H contains uncountably many non-local separating points leads to a contradiction.

5. In conclusion we point out that the property of being a hereditary arc sum is invariant under continuous transformations which are either *monotone*[7] or *interior*.[8]

[5] G. T. Whyburn, *Decompositions of continua by means of local separating points*, American Journal of Mathematics, vol. 55(1933), pp. 437–457. See Lemma 2.5.

[6] Loc. cit., p. 452.

[7] G. T. Whyburn, *Non-alternating transformations*, American Journal of Mathematics, vol. 56(1934), pp. 294–302, Theorem 1.3.

[8] G. T. Whyburn, *Interior transformations on compact sets*, this Journal, vol. 3(1937), pp. 370–381.

THEOREM F. *Let H be a hereditary arc sum. If $B = T(H)$ is the continuous image of H under T which is either monotone or interior, then B is a hereditary arc sum.*

Proof. Let T be monotone. Since H is compact, the inverse of a connected set in B is connected. Let S be any open connected set in B. It suffices to prove that the boundary of S is countable. The set $R = T^{-1}(S)$ is open and connected. Since H is a hereditary arc sum, $B(R)$ is countable. Now $\overline{T(R)} - T(R) \subset T(\bar{R} - R)$. Thus $\bar{S} - S = \overline{T(R)} - T(R)$ is countable.

Let T be an interior transformation. We need the following results regarding such transformations:[8]

(i) If A is compact, R open in A and T interior on A, then $B(T(R)) \subset T(B(R))$.

(ii) Let R be an open connected set in $B = T(A)$, where A is compact and T interior on A; then each quasi-component Q of $T^{-1}(R)$ maps onto all of R under T.

Suppose that A is a hereditary arc sum. Let R be any open connected subset of B. The set $S = T^{-1}(R)$ is open in the compact locally connected continuum A. Hence each component of S is open in A. For locally connected sets the components and quasi-components are identical. Thus each component Q of S maps onto all of R under T. Using (i), we have $B(T(Q)) = B(R) \subset T(B(Q))$. Since Q is a region in the hereditary arc sum A, $B(Q)$ is countable, and hence $B(R)$ is likewise.

OREGON STATE COLLEGE.

A GENERALIZATION OF PICARD'S AND RELATED THEOREMS

By Raphael M. Robinson

With the exception of the generalization of Picard's second theorem, the results contained in this paper are already known. But new (and quite simple) proofs of the various theorems to which the title refers are given.[1] The necessary preliminary material has also been included, so that only the elements of function theory are assumed as known. In particular, the theorems with whose generalizations we are concerned are not assumed, but are obtained as special cases of our results.

1. **Introduction.** An analytic function $f(z)$ is said to be meromorphic in a given region if it has no singularities there except poles. And indeed, in considering meromorphic functions, a pole is not to be regarded as a singularity, but rather simply as a point where the value of the function is ∞. The function which is equal to ∞ throughout the given region is also to be regarded as meromorphic there. By the expression "meromorphic function" we shall mean a function which is meromorphic for all values of z except $z = \infty$. (Since a function which is meromorphic for all values of z is a rational function, no new term is needed in that case.)

We shall call a number a an excluded value for $f(z)$ if the equation $f(z) = a$ has no root in the given region. Picard's first theorem may then be stated as follows: *A meromorphic function which admits more than two excluded values is a constant.*

We shall say that a is an exceptional value of order m if the equation $f(z) = a$ has no root of multiplicity less than m. The symbol m in that statement represents a positive integer or ∞. For $m = 1$ no condition is imposed on the function, while for $m = \infty$ the statement says that $f(z) = a$ has no root of any order. We wish now to assign weights to the exceptional values; we shall say, so to speak, that an exceptional value of order m has an importance which is a certain fraction of that of an excluded value. From what we have said, it is clear that the weight should increase with m and that it should be 0 for $m = 1$ and 1 for $m = \infty$. The expression $1 - 1/m$ suggests itself for the weight; we shall take this as the definition, since it proves useful. (This idea of weight

[1] The previously known proofs depend on Nevanlinna's theory of meromorphic functions; see his book, *Le Théorème de Picard-Borel*, 1929, p. 102. For a discussion of the results from a different point of view, see Ahlfors, *Quelques propriétés des surfaces de Riemann*, Bull. de la Soc. Math. de France, vol. 60(1932), pp. 197–207. The generalized criterion for a normal family of meromorphic functions seems to have been stated first by Bloch (making use of Nevanlinna's results); see, e.g., Valiron, *Familles Normales de Fonctions Méromorphes*, 1929, p. 21.

was used by Valiron, loc. cit.) By the total weight of a number of exceptional values, we mean the sum of their weights. With this terminology, the generalization of Picard's theorem with which we are concerned may be stated as follows: *A meromorphic function which admits exceptional values of total weight more than 2 is a constant.*

In this paper, we give a proof of this theorem, as well as of the corresponding generalizations of the theorems of Landau, Schottky, and Montel (criterion for a normal family of meromorphic functions), and of Picard's second theorem. In the last section, it is shown that none of these theorems can be replaced by better theorems of the same type, i.e., theorems whose hypotheses are concerned solely with the exceptional values and their orders.

2. **The five principal cases.** Before proceeding to the proof of our theorems, we determine the cases in which the total weight of the exceptional values is more than 2. We find that there are essentially only five cases; that is, that all others are included in these in the sense of having exceptional values of higher order (or greater weight) than the corresponding exceptional values of one of the five principal cases. The principal cases are those in which the orders of the exceptional values are

$$(2, 2, 2, 2, 2), \quad (2, 2, 2, 3), \quad (2, 3, 7), \quad (2, 4, 5), \quad (3, 3, 4).$$

This may be seen as follows. We may suppose that the order of each exceptional point is at least 2, and hence that its weight is at least $\frac{1}{2}$. Thus if there are five or more exceptional values, the total weight is more than 2; and such a case is included in the first case above. If there are four exceptional values, one of them must be of order more than 2; such a case is included in the second case above. If there are three exceptional values, of orders m_1, m_2, m_3, we must have

$$\frac{1}{m_1} + \frac{1}{m_2} + \frac{1}{m_3} < 1.$$

We may suppose the three values arranged in such an order that $m_1 \leqq m_2 \leqq m_3$. It is then easily seen that one of the following sets of inequalities must hold: $m_1 \geqq 2$, $m_2 \geqq 3$, $m_3 \geqq 7$; $m_1 \geqq 2$, $m_2 \geqq 4$, $m_3 \geqq 5$; $m_1 \geqq 3$, $m_2 \geqq 3$, $m_3 \geqq 4$. We are thus led to the last three of the five principal cases. It is impossible to have fewer than three exceptional values, since each weight is at most 1, and the total weight is to be more than 2. Thus the five cases listed are essentially the only ones.

If we suppose that the function being considered has an excluded value (as, for example, when we consider functions regular in a given region), the five cases reduce to only two; namely, those in which the orders of the remaining exceptional values are either (2, 2, 2) or (2, 3).

To illustrate the meaning of these cases for the theorems which we shall consider, we formulate the generalization of Picard's first theorem, stated in §1,

for the special case of entire functions, in the following form. *Let $f(z)$ be an entire function which is not constant. Then if w_1, w_2, w_3 are any complex numbers, one of the equations $f(z) = w_k$ $(k = 1, 2, 3)$ has a simple root; and if w_1 and w_2 are any complex numbers, either the equation $f(z) = w_1$ has a simple root, or the equation $f(z) = w_2$ has a simple or double root.* Picard's theorem (for entire functions) is evidently included in the second part of this statement.

3. **Families of meromorphic functions.** In the study of meromorphic functions, it is convenient to consider, in place of the plane of the dependent variable, its stereographic projection on a unit sphere about the origin, the center of projection being the point whose space coördinates are (0, 0, 1). We shall denote by (w', w'') the spherical distance between the points w' and w'', i.e., the length of the shorter arc of the great circle joining them.

We shall say that a sequence $f_n(z)$ of meromorphic functions converges uniformly on a given set of points to a meromorphic function $f(z)$ if to every $\epsilon > 0$ there is an N such that $(f(z), f_n(z)) < \epsilon$ for $n \geqq N$ and for every z in the given point set. That is, the uniformity of the convergence is to be measured on the sphere.

Let us now consider a family of functions meromorphic in $|z| < 1$. (The results may easily be extended to a family of functions meromorphic in any given region.)

We shall say that the functions are equicontinuous in $|z| \leqq r$ (where $r < 1$) if to every $\epsilon > 0$ there is a $\delta > 0$ such that $(f(z'), f(z'')) < \epsilon$ for $|z' - z''| < \delta$, for every function $f(z)$ of the family and every pair of points z' and z'' in the circle $|z| \leqq r$. We shall say that the functions are equicontinuous in the interior of the unit circle if they are equicontinuous in $|z| \leqq r$ for each $r < 1$.

We shall say that the functions form a normal family in the unit circle if from any sequence of them a subsequence can be chosen which, for any $r < 1$, converges uniformly for $|z| \leqq r$.

With these definitions, we may state the following fundamental theorem, which we may designate as Ostrowski's form of Arzelà's theorem:[2] *A family of functions meromorphic in the unit circle form a normal family if and only if they are equicontinuous in the interior of the unit circle.*

The theorem may be proved as follows. Suppose first that the functions are equicontinuous in the interior of the unit circle. Let $z_1, z_2, z_3, \cdots$ be the points within the unit circle having rational coördinates, arranged into a sequence in any manner. Consider any sequence of functions from our family. We choose a subsequence $f_{1n}(z)$ which converges at z_1; from this we choose a subsequence $f_{2n}(z)$ which converges also at z_2; and so forth. The sequence $f_{nn}(z)$ will then converge at all the rational points within the unit circle. We shall show that this sequence converges uniformly for $|z| \leqq r$, if $r < 1$. If $\epsilon > 0$ is given, we

[2] The corresponding theorem for bounded real functions was stated by Arzelà in 1894; a special case had been stated even earlier by Ascoli. Ostrowski stated the result for meromorphic functions in Math. Zeitschrift, vol. 24(1925), p. 233.

can choose $\delta > 0$, so that $(f_{nn}(z'), f_{nn}(z'')) < \frac{1}{3}\epsilon$ for $|z' - z''| < \delta$, for all values of n, and for any pair of points z' and z'' in $|z| \leqq r$. We then choose a finite number of rational points in $|z| \leqq r$, such that there is one of them within a distance less than δ from each point of that circle. If we choose N so large that $(f_{mn}(z), f_{nn}(z)) < \frac{1}{3}\epsilon$, for $m \geqq N$, $n \geqq N$ and z one of that finite set of points, then $(f_{mn}(z), f_{nn}(z)) < \epsilon$ throughout $|z| \leqq r$, for those values of m and n. Hence the sequence $f_{nn}(z)$ converges uniformly in $|z| \leqq r$. The functions thus form a normal family.

Conversely, we suppose that the functions form a normal family. We have then to prove that they are equicontinuous in $|z| \leqq r$, for each $r < 1$. Suppose that this is not the case. Then for some $r < 1$ and for some $\epsilon > 0$, there is no $\delta > 0$ such that $(f(z'), f(z'')) < \epsilon$ for $|z' - z''| < \delta$, for each pair of points z' and z'' in $|z| \leqq r$, and for each function $f(z)$ of the family. For this r and ϵ, therefore, we can find a sequence $f_n(z)$ of functions of the family, and a sequence of pairs of points z_n' and z_n'' in $|z| \leqq r$ such that $(f_n(z_n'), f_n(z_n'')) \geqq \epsilon$ and $|z_n' - z_n''| \to 0$ for $n \to \infty$. We may suppose that the numbers z_n' converge to a limit z_∞, since otherwise a subsequence could be chosen which would converge. Then z_n'' converges to z_∞ also. Since the functions form a normal family, we may suppose in the same way that the sequence of functions $f_n(z)$ converges uniformly for $|z| \leqq r$ to a meromorphic function $f_\infty(z)$. Then each of the distances $(f_\infty(z_n'), f_n(z_n'))$, $(f_\infty(z_n''), f_n(z_n''))$, and $(f_\infty(z_n'), f_\infty(z_n''))$ can be made less than $\frac{1}{3}\epsilon$ by taking n sufficiently large. Thus we have $(f_n(z_n'), f_n(z_n'')) < \epsilon$ for those values of n. The hypothesis that the functions are not equicontinuous in $|z| \leqq r$ leads therefore to a contradiction; hence the functions are equicontinuous in $|z| \leqq r$, as was to be proved.

4. **Curvature of a metric; special metrics.** We shall have occasion to make use of several different definitions of the length of a curve and of distance, or, as one says, of different metrics. In each case an element of length ds will be defined in the form $\lambda(z)\,|dz|$, where $\lambda(z) \geqq 0$ is continuous. (That we use this special form of metric means that we use an equal scale in making measurements in any direction from a given point.) The length of a curve is then obtained by integrating ds along the curve. The distance between two points is defined as the length of the shortest curve joining them (supposing such a curve to exist).

If $\lambda(z) > 0$ and has continuous second partial derivatives with respect to x and y (where $z = x + iy$), we may speak of the curvature of the metric $ds = \lambda(z)\,|dz|$. By the curvature of this metric we understand[3] the quantity

$$-\frac{\Delta \log \lambda}{\lambda^2},$$

where Δ is the Laplacian operator, i.e., $\Delta f(x, y) = f_{xx}(x, y) + f_{yy}(x, y)$.

The curvature of a metric is invariant under a conformal transformation.

[3] That this definition of curvature agrees with that given in differential geometry for the Gaussian curvature of a surface in space realizing our metric is for our purposes inessential.

More exactly, if z and z^* are analytic functions of each other which are regular at the points under consideration, if $\lambda^*(z^*)$ is determined from $\lambda^*(z^*) \mid dz^* \mid = \lambda(z) \mid dz \mid$, and if Δ^* is the Laplacian operator in the variables x^* and y^* ($z^* = x^* + iy^*$), then

$$\frac{\Delta^* \log \lambda^*}{\lambda^{*2}} = \frac{\Delta \log \lambda}{\lambda^2}.$$

To verify this, one may first establish the manner in which the Laplacian operator transforms; namely, according to the equation $\Delta^* f^* \cdot \mid dz^* \mid^2 = \Delta f \cdot \mid dz \mid^2$, where $f(x, y) = f^*(x^*, y^*)$ is a function with continuous second partial derivatives. In showing this, use is made of the fact that the real and imaginary parts of an analytic function satisfy the Cauchy-Riemann and Laplace equations. Finally, we apply this formula, again making use of the fact that the real part of an analytic function satisfies the Laplace equation, to obtain the desired result.

We now consider some special metrics. If $\lambda(z)$ is a positive constant, we have a Euclidean metric. In this case, the curvature is 0. Aside from this case, we have already considered in the last section the spherical metric, though not from the point of view of first defining an element of length. But since the stereographic projection is a conformal transformation, the element of length on the sphere would be of the required sort. We shall denote this element of length by $d(z)$. (It is not necessary for us to have an explicit formula for $d(z)$. However, this is readily calculated, namely,

$$d(z) = \frac{2 \mid dz \mid}{1 + \mid z \mid^2}.$$

Using this formula, we can calculate the curvature of the spherical metric and find that it is equal to 1.)

Another classical metric, which applies, however, only to the upper half-plane $y > 0$, is that in which an element of length, which we shall denote by $d[z]$, is defined by the equation

$$d[z] = \frac{\mid dz \mid}{y}.$$

This determination of distance is especially significant because it is invariant under a linear (fractional) transformation of the upper half-plane into itself. (Such a transformation has the form $z^* = (az + b)/(cz + d)$, with real a, b, c, d and $ad - bc > 0$. It is then simple to calculate the invariance of $\mid dz \mid/y$.) We find that the curvature of this metric is -1.

In this metric, the shortest curve joining two points is an arc of a circle orthogonal to the real axis. For by means of a linear transformation of the upper half-plane into itself, which leaves $d[z]$ invariant, we can transform the two points into two points on a vertical line. In this case, it is easy to see that the line segment joining them is the shortest such curve, according to this metric.

If we map the upper half-plane onto the interior C of any circle by means of a linear transformation, we introduce a metric with similar properties in C. The element of length is invariant under a linear transformation of C into itself, the curvature is -1, and the shortest curves are circles orthogonal to the circumference of C. If we call these orthogonal circles "straight lines", then the geometry in C with this metric is hyperbolic geometry. Consequently, we shall speak of this metric as the hyperbolic metric of C. We shall use the notation $d[z]$ for an element of length in this case also, and the notation $[z', z'']$ for the distance between the points z' and z''. In using this notation, we must specify what C is. Thus $d[z] = |dz|/y$ if C is the upper half-plane, while $d[z] = |dz|/x$ if C is the right half-plane. (Another interesting case is that in which C is the unit circle. We do not need the explicit formula in that case, but it is easily calculated, namely,

$$d[z] = \frac{2|dz|}{1-|z|^2}.$$

It may be noticed that this is very similar in form to the expression for $d(z)$.)

5. **Ahlfors's generalization of Schwarz's lemma.** If $f(z)$ is regular for $|z| < 1$, if $|f(z)| < 1$ there, and if $f(0) = 0$, then, according to Schwarz's lemma, we must have $|f'(0)| \leqq 1$. If we suppose only that $f(z)$ is regular for $|z| < 1$ and $|f(z)| < 1$ there, then putting $w = f(z)$, and making use of the hyperbolic metric in the unit circle in both the z- and w-planes, we can say that $d[w] \leqq d[z]$. For by making linear transformations in both the z- and w-planes which leave the unit circle invariant, but take the points z and w under consideration to the origin in their respective planes, we can reduce this to the previous statement.

Ahlfors's generalization[4] of this consists in considering any function $f(z)$ meromorphic in $|z| < 1$ and measuring distances on the Riemann surface traced out by $w = f(z)$ according to any metric ds whose curvature is $\leqq -1$, the result obtained being that $ds \leqq d[z]$. The conditions to be imposed on the metric at the branch points and points at infinity are considered below. (Evidently the same result would be true if we considered a function $f(z)$ meromorphic in any circle C, if $d[z]$ denoted the hyperbolic element of distance for the circle C.)

We consider first the following simpler case. Suppose that we have defined in the unit circle, besides the hyperbolic metric, another metric $ds = \lambda(z)|dz|$, whose curvature is $\leqq -1$ at all those points where $\lambda(z) > 0$. We shall show that $ds \leqq d[z]$. Let $d[z] = \lambda_1(z)|dz|$ (i.e., $\lambda_1(z) = 2/(1-|z|^2)$); we have to show that $\lambda(z) \leqq \lambda_1(z)$ for $|z| < 1$. Let $\lambda_r(z)|dz|$ be the hyperbolic element of distance for the circle $|z| < r$. Then it is sufficient to show that, for each $r < 1$, $\lambda(z) \leqq \lambda_r(z)$ for $|z| < r$. Now $\Delta \log \lambda_r = \lambda_r^2$; and $\Delta \log \lambda \geqq \lambda^2$ at those points where $\lambda > 0$. Hence, if $\lambda > 0$,

$$\Delta \log \frac{\lambda}{\lambda_r} \geqq \lambda^2 - \lambda_r^2.$$

[4] Ahlfors, *An extension of Schwarz's lemma*, Trans. Amer. Math. Soc., vol. 43(1938), pp. 359–364.

Now at any point where $\lambda > \lambda_r$, we have $\lambda > 0$, so that we can use the above inequality to obtain $\Delta \log (\lambda/\lambda_r) > 0$. Therefore, at such a point, $\log (\lambda/\lambda_r)$ cannot have a maximum. For if a function $f(x, y)$ with second partial derivatives has a maximum at a point (x, y), then $f_{xx}(x, y) \leqq 0$ and $f_{yy}(x, y) \leqq 0$, so that also $\Delta f(x, y) \leqq 0$. Hence λ/λ_r cannot have a maximum which is >1 at any point in $|z| < r$. Now $\lambda_r(z) \to \infty$ for $|z| \to r$, while $\lambda(z)$ remains bounded, since it is continuous in $|z| < 1$. Hence $\lambda/\lambda_r \to 0$ for $|z| \to r$. Combining this with the fact that λ/λ_r has no maximum > 1 in $|z| < r$, we see that $\lambda/\lambda_r \leqq 1$ there, as was to be shown.

If now $w = f(z)$ maps $|z| < 1$ on a Riemann surface over the w-plane, and we define on that surface a metric whose curvature is $\leqq -1$, that metric may be transferred back to $|z| < 1$, and the curvature does not change. We must suppose that the metric is such near the branch points and points at infinity on the surface that, when a neighborhood of such a point is mapped onto a plane finite region, the metric is of the kind we have considered. That is, the transformed λ, say λ^*, must either be 0 at the point corresponding to the branch point or point at infinity, or have continuous second partial derivatives there. That an element of length on the Riemann surface, in the metric defined there, is less than the corresponding $d[z]$ is Ahlfors's result.

We must now examine more closely the condition concerning the branch points and points at infinity. For our purposes it will be sufficient to consider only the case in which $\lambda^* = 0$ at the corresponding point, after the mapping mentioned in the last paragraph. Now such a mapping is done by the function $w^* = (w - a)^{1/m}$, in case of a branch point with m sheets at a finite point a. Then

$$dw^* = \frac{1}{m}(w - a)^{(1/m)-1} dw,$$

so that if $ds = \lambda(w)\,|dw| = \lambda^*(w^*)\,|dw^*|$, we have

$$\lambda^* = \lambda \left| \frac{dw}{dw^*} \right| = \lambda m\, |w - a|^{1-1/m}.$$

Hence the condition imposed becomes

$$|w - a|^{1-1/m} \cdot \lambda \to 0 \qquad \text{for } w \to a.$$

This condition may also be written, using the spherical metric, in the form

$$(w, a)^{1-1/m} \frac{ds}{d(w)} \to 0 \qquad \text{for } w \to a,$$

and in this form it is valid even if $a = \infty$.

6. **Construction of a metric.** The purpose of this section is to construct a metric with curvature $\leqq -1$, defined over the entire sphere, with the exception of a finite number of points; at these points it is to have singularities of a prescribed sort.

LEMMA.[5] *Given the real numbers* $\gamma_1, \gamma_2, \cdots, \gamma_n, \gamma$ *such that*

$$\gamma_k < 1 \ (1 \leqq k \leqq n), \qquad \gamma < -1, \qquad \gamma_1 + \cdots + \gamma_n + \gamma > 0.$$

Then there exists a function $\lambda(w) > 0$ *with continuous second partial derivatives, defined over the entire w-plane, except at the points* $w_1, w_2, \cdots, w_n, \infty$, *which satisfies the inequality*

$$\Delta \log \lambda(w) \geqq \lambda(w)^2,$$

and the boundary conditions

$$\lambda(w) \sim \text{const}/|\, w - w_k \,|^{\gamma_k} \qquad \text{for } w \to w_k,$$

$$\lambda(w) \sim \text{const} \cdot |\, w \,|^{\gamma} \qquad \text{for } w \to \infty.$$

Proof. Let $r_k = |\, w - w_k \,|$ and $r = |\, w \,|$. Let ϵ denote any sufficiently small positive number, and let $c > 0$ be determined from $nc\epsilon = \gamma_1 + \cdots + \gamma_n + \gamma$. Then put

$$\mu(w) = \prod_{k=1}^{n} \frac{(1 + r_k^{\epsilon})^c}{r_k^{\gamma_k}},$$

and finally

$$\lambda(w) = \beta \cdot \mu(w),$$

where β is a sufficiently small positive constant. We shall verify that by proper choice of ϵ and β the conditions of the lemma are satisfied by this $\lambda(w)$.

We see at once that

$$\mu \sim \text{const}/r_k^{\gamma_k} \qquad \text{for } w \to w_k,$$

and, using the definition of c, that

$$\mu \sim r^{\gamma} \qquad \text{for } w \to \infty.$$

These equations show that λ satisfies the required boundary conditions.

If we apply the general formula

$$\Delta f(r) = f''(r) + \frac{1}{r} f'(r)$$

to the function $\log (1 + r^{\epsilon})$, we find that

$$\Delta \log (1 + r^{\epsilon}) = \frac{\epsilon^2 r^{\epsilon-2}}{(1 + r^{\epsilon})^2}.$$

Using this result and the fact that $\Delta \log r = 0$, both of which remain true if r is replaced by r_k, we obtain the formula

$$\Delta \log \mu = c\epsilon^2 \sum_{k=1}^{n} \frac{r_k^{\epsilon-2}}{(1 + r_k^{\epsilon})^2}.$$

[5] This lemma is a special case of a theorem due to Picard; namely, that under the same hypotheses, we can satisfy an equality rather than an inequality. See his paper, *De l'intégration de l'équation* $\Delta u = e^u$ *sur une surface de Riemann fermée*, Journal für Math., vol. 130(1906), pp. 243–258, as well as earlier papers mentioned there.

From this we find that

$$\Delta \log \mu \sim c\epsilon^2 r_k^{\epsilon-2} \qquad \text{for } w \to w_k,$$

and that

$$\Delta \log \mu \sim nc\epsilon^2 r^{-\epsilon-2} \qquad \text{for } w \to \infty.$$

Therefore

$$\frac{\Delta \log \mu}{\mu^2} \sim \frac{\text{const}}{r_k^{2-2\gamma_k-\epsilon}} \qquad \text{for } w \to w_k$$

(the constant being positive), and

$$\frac{\Delta \log \mu}{\mu^2} \sim nc\epsilon^2 r^{-2-2\gamma-\epsilon} \qquad \text{for } w \to \infty.$$

Hence, if ϵ is sufficiently small (remembering that $2 - 2\gamma_k$ and $-2 - 2\gamma$ are positive), we have

$$\frac{\Delta \log \mu}{\mu^2} \to +\infty \qquad \text{for } w \to w_k \text{ and } w \to \infty.$$

Hence $(\Delta \log \mu)/\mu^2$ has a positive minimum. Now

$$\frac{\Delta \log \lambda}{\lambda^2} = \frac{\Delta \log \mu}{\beta^2 \mu^2},$$

so that by taking β sufficiently small, we can insure that, for all values of w except $w_1, w_2, \cdots, w_n, \infty$, we have the inequality

$$\frac{\Delta \log \lambda}{\lambda^2} \geqq 1.$$

For later purposes, we need to know the behavior of the function λ determined in this lemma, when the $w_1, w_2, \cdots, w_n$ are variable. Now λ was given by an explicit formula, except that undetermined constants ϵ, c, and β were involved. The constant ϵ may be chosen independent of $w_1, \cdots, w_n$, since it was subject only to the conditions $\epsilon < 2 - 2\gamma_k$ for $1 \leqq k \leqq n$, and $\epsilon < -2 - 2\gamma$. The constant c is then also independent of the w_k. Finally, the constant β may be chosen as a continuous function of $w_1, \cdots, w_n$, these n points being supposed distinct from each other and from ∞. To verify this, we note first that all of the limit relations stated in the proof of the lemma for $w \to w_k$ or $w \to \infty$ are true under the more general hypothesis that the points $w_1, \cdots, w_n$ are variable, but that $w \to w_k'$ or $w \to \infty$, and that $w_l \to w_l'$ for $1 \leqq l \leqq n$, where $w_1', \cdots, w_n'$ are a fixed set of values of $w_1, \cdots, w_n$. The function $(\Delta \log \mu)/\mu^2$ is a continuous function of $w, w_1, \cdots, w_n$, as long as w is different from the singular points, and approaches $+\infty$ under the conditions just mentioned. That is, it is continuous even if w is equal to some w_k or ∞, if we define it as $+\infty$ at those points, and understand continuity in

an extended sense which does not exclude the function having an infinite value. Now β was subject only to the condition that $\Delta \log \mu \geqq \beta^2 \mu^2$; for given values of the w_k, we may choose β as the square root of the minimum, for all values of w, of $(\Delta \log \mu)/\mu^2$. With this choice, β is a continuous function of $w_1, \cdots, w_n$. Then λ itself is a continuous function of $w, w_1, \cdots, w_n$, in the extended sense just mentioned, if we suppose, as always, that $w_1, \cdots, w_n, \infty$ are distinct. We note also that $\lambda \cdot |w|^2$ is continuous, in the same sense, for $w = \infty$.

THEOREM. *Given the points $w_1, w_2, \cdots, w_q$ on the sphere, and the real numbers $\alpha_1, \alpha_2, \cdots, \alpha_q$ such that $0 < \alpha_k \leqq 1$ for $1 \leqq k \leqq q$, but $\alpha_1 + \alpha_2 + \cdots + \alpha_q > 2$. Then it is possible to find a metric ds whose curvature is $\leqq -1$, defined over the entire sphere, except at the points $w_1, \cdots, w_q$, and such that $ds/d(w)$ becomes infinite in such a way at these points that*

$$(w, w_k)^{\alpha_k} \frac{ds}{d(w)} \to 0 \qquad \text{for } w \to w_k \ (1 \leqq k \leqq q).$$

The metric can be determined so that it varies continuously with $w_1 \cdots, w_q$.

Proof. Suppose first that $w_q = \infty$, and apply the lemma with $n = q - 1$. The γ's are to be so chosen that $0 < \gamma_k < \alpha_k$ for $1 \leqq k \leqq n$, while $-2 < \gamma < -2 + \alpha_q$; we see at once that it is possible to do this in such a way that the conditions of the lemma are satisfied. We determine $\lambda(w)$ from the lemma and put $ds = \lambda(w) \, |dw|$; this metric has a curvature which is $\leqq -1$. We find also that

$$\frac{ds}{d(w)} \sim \text{const} \cdot \lambda(w) \to \infty \qquad \text{for } w \to w_k \ (1 \leqq k \leqq n),$$

and

$$\frac{ds}{d(w)} \sim \text{const} \cdot \lambda(w) \, |w|^2 \to \infty \qquad \text{for } w \to \infty.$$

(To derive the last formula, we must make use of the fact that $d(w) \sim \text{const} \cdot |dw|/|w|^2$ for $w \to \infty$. This may be seen at once from the formula for $d(w)$, but may also be derived without using that formula in the following way. Let $w^* = 1/w$; then $d(w^*) = d(w)$ and $|dw^*| = |dw|/|w|^2$. Since $d(w^*) \sim \text{const} \cdot |dw^*|$ for $w^* \to 0$, we have the required result.) On the other hand, we find that

$$(w, w_k)^{\alpha_k} \frac{ds}{d(w)} \sim \text{const} \cdot |w - w_k|^{\alpha_k} \lambda(w) \to 0 \qquad \text{for } w \to w_k \ (1 \leqq k \leqq n),$$

$$(w, \infty)^{\alpha_q} \frac{ds}{d(w)} \sim \text{const} \cdot |w|^{-\alpha_q} \lambda(w) |w|^2 \to 0 \qquad \text{for } w \to \infty.$$

Thus all the conditions are satisfied, and the metric varies continuously with $w_1, \cdots, w_n$, if we suppose that the γ's were chosen to depend only on the α's.

More precisely, $ds/d(w)$ is a continuous function (in the extended sense) of w, $w_1, \cdots, w_n$; for finite values of w this follows from the continuity of λ $(= ds/|\,dw\,|)$, and for $w = \infty$ from the continuity of $\lambda \cdot |\,w\,|^2$ $(\sim \text{const} \cdot ds/d(w))$.

Finally, for the general case, where we do not suppose that $w_q = \infty$, we can, by a rotation of the sphere, transform to the first case, and indeed the rotation can be made to depend continuously on w_q, so that the condition about continuous variation is also satisfied.

7. **The generalizations of Picard's and related theorems.** We have now covered all the necessary preliminaries and come to the proofs of the generalizations of Picard's and related theorems. After each of our theorems, we state one of the classical theorems, which follows easily from it.

The letters K, $K_1, \cdots$ will denote positive constants, which depend only on the quantities indicated. The letters δ and r denote positive numbers, and $r < 1$. If we put $w = f(z)$, then w' will denote $f(z')$, and similarly in other cases.

THEOREM 1. *Let $f(z)$ be meromorphic in a circle C and admit exceptional values of total weight more than 2, the spherical distance between any two exceptional values being $\geqq \delta$. Then if we put $w = f(z)$, we have*

$$d(w) \leqq K(\delta)d[z],$$

and consequently also

$$(w', w'') \leqq K(\delta)[z', z''].$$

(The symbols $d[z]$ and $[z', z'']$ refer to the hyperbolic metric in C.)

Proof. If we suppose that the exceptional values are $w_1, w_2, \cdots, w_q$, with orders $m_1, m_2, \cdots, m_q$, and if we apply the theorem of §6 with $\alpha_k = 1 - 1/m_k$, we obtain a metric ds with curvature $\leqq -1$, defined over the entire sphere except at the points $w_1, \cdots, w_q$. We wish to use this metric on each sheet of the Riemann surface traced out by $w = f(z)$. Now any value of z, which is a root of $f(z) = w_k$ of multiplicity m, corresponds to a branch point with m sheets at w_k, on the Riemann surface. By hypothesis, this m is at least m_k. Consequently, the condition at the end of §5 is satisfied, so that we can apply Ahlfors's theorem, thus obtaining the result that $ds \leqq d[z]$.

The ratio $ds/d(w)$ is always positive and is a continuous function (in the extended sense) of w, $w_1, \cdots, w_q$, as long as $w_1, \cdots, w_q$ are distinct. Hence if we restrict the arguments to a closed set by means of the conditions $(w_k, w_l) \geqq \delta$ (for $1 \leqq k < l \leqq q$), then $ds/d(w)$ has a positive minimum; hence

$$d(w) \leqq K(\delta; m_1, \cdots, m_q)ds.$$

The arguments $m_1, \cdots, m_q$ may be struck out, since, by §2, there are only a finite number of cases that need to be considered. Therefore

$$d(w) \leqq K(\delta)ds \leqq K(\delta)d[z].$$

LANDAU'S THEOREM. *If $f(z)$ is regular for $|z| < 1$ and is never equal to 0 or 1 there, then*

$$|f'(0)| \leqq K_1(f(0)).$$

THEOREM 2. *A meromorphic function which admits exceptional values of total weight more than 2 is a constant.*

Proof. If $f(z)$ is the function, then by Theorem 1

$$(f(z'), f(z'')) \leqq K(\delta)[z', z''],$$

where $[z', z'']$ is the distance between the points z' and z'' in the hyperbolic metric of any circle C containing those points. If we take C as $|z| < R$ and let $R \to \infty$, keeping z' and z'' fixed, then $[z', z''] \to 0$. Consequently, $f(z') = f(z'')$; that is, $f(z)$ is constant.

PICARD'S FIRST THEOREM. *An entire function which is never equal to 0 or 1 is a constant.*

THEOREM 3. *The functions meromorphic in $|z| < 1$, which admit exceptional values of total weight >2 and minimum spherical distance $\geqq\delta$, form a normal family.*

Proof. From Theorem 1, we obtain at once

$$(f(z'), f(z'')) \leqq K_2(\delta, r)\,|z' - z''|,$$

for any pair z' and z'' in the circle $|z| \leqq r$. Thus the functions are equicontinuous in the interior of the unit circle, and hence, by §3, form a normal family.

MONTEL'S THEOREM. *The functions which are meromorphic in $|z| < 1$ and admit there three given excluded values form a normal family.*

THEOREM 4. *Suppose that $f(z)$ is regular for $|z| < 1$ and admits finite exceptional values of total weight > 1, whose spherical distance from each other and from ∞ is $\geqq\delta$. If $|f(0)| \leqq M$, then*

$$|f(z)| \leqq K_3(r, \delta, M) \qquad \text{for } |z| \leqq r.$$

Proof. Otherwise, for some values of r, δ, M, we could find a sequence of functions $f_n(z)$, satisfying the conditions of the theorem, and a sequence of points z_n in the circle $|z| \leqq r$ and such that $f_n(z_n) \to \infty$. We may suppose that the points z_n converge to a limit z_∞, since otherwise we could choose a convergent subsequence. Let R be chosen between r and 1. Then, since the functions considered form a normal family, we may suppose that the sequence $f_n(z)$ converges uniformly (in the spherical sense) in $|z| \leqq R$, to a meromorphic function $f_\infty(z)$; for otherwise an appropriate subsequence could be chosen. We must have $f_\infty(z_\infty) = \infty$. On the other hand, $|f_\infty(0)| \leqq M$, so that $f_\infty(z)$ is not identically ∞, but has a pole at z_∞. But a sequence of functions regular in $|z| < R$ cannot converge uniformly toward a function having a pole there, as may be seen, e.g., by using the principle of maximum. The hypothesis that the theorem is false thus leads to a contradiction.

SCHOTTKY'S THEOREM. *If $f(z)$ is regular in $|z| < 1$ and always different from 0 and 1 there, then*

$$|f(z)| \leqq K_4(r, f(0)) \qquad \text{for } |z| \leqq r.$$

We come now to the generalization of Picard's second theorem, which is believed to be new. It may be pointed out that the proof necessarily proceeds along different lines from the proof commonly given for Picard's second theorem itself; for that proof makes use of the principle of maximum, which is not available here, since we are considering meromorphic functions.

THEOREM 5. *A function $f(z)$ which is single-valued and meromorphic in $1 < |z| < \infty$ and admits exceptional values of total weight >2 cannot have an essential singular point at $z = \infty$.*

Proof. Suppose that ∞ is an essential singularity for $f(z)$. Then neither $f(z)$ nor $1/f(z)$ can be bounded near $z = \infty$. Hence we can find a point z', with $|z'| > \max(e^{6K}, e^{K^2})$ (where $K = K(\delta)$ is the constant of Theorem 1), and such that $(f(z'), \infty) < \frac{1}{6}\pi$. A second point z'' may be found, with $|z''| > |z'|$, and such that $(f(z''), 0) < \frac{1}{6}\pi$.

Let $\zeta = \log z$ and $g(\zeta) = f(e^{\zeta})$, so that $f(z) = g(\zeta)$. Then $g(\zeta)$ is meromorphic for $\xi > 0$, where $\zeta = \xi + i\eta$. Furthermore, $g(\zeta) = a$ has roots of the same order as $f(z) = a$, so that $g(\zeta)$ admits exceptional values of total weight >2. Thus we shall be able to apply Theorem 1. Let ζ be any point on the line segment joining the points $\zeta' - \pi i$ and $\zeta' + \pi i$. Then

$$(g(\zeta), g(\zeta')) \leqq K \cdot [\zeta, \zeta'] \leqq K \int_{\zeta'}^{\zeta} \frac{|d\zeta|}{\xi},$$

where the integral is taken along the line segment joining ζ' to ζ. Now $\xi > 6K$ there, and the length of the segment (in the ordinary sense) is $\leqq \pi$. Hence $(g(\zeta), g(\zeta')) < \frac{1}{6}\pi$ for ζ on the given segment; that is, $(f(z), f(z')) < \frac{1}{6}\pi$ for $|z| = |z'|$. From this, we conclude that $(f(z), \infty) < \frac{1}{3}\pi$ for $|z| = |z'|$. In exactly the same way, we find that $(f(z), 0) < \frac{1}{3}\pi$ for $|z| = |z''|$.

Thus the circular ring $|z'| \leqq |z| \leqq |z''|$ is mapped by $f(z)$ on a part of the sphere, one of whose boundaries is within a distance $\frac{1}{3}\pi$ from ∞ (i.e., within 60° from the north pole), and whose other boundary is within a distance $\frac{1}{3}\pi$ from 0 (i.e., within 60° from the south pole). Consequently, the map of the circular ring must cover the zone extending 30° on either side of the equator. (The zone may be covered multiply.) Hence the area of the map must be at least 2π, since that is the area of this zone.

On the other hand, the area of that map is not more than K^2 times the hyperbolic area of the corresponding region in the ζ-plane. Take this region as

$$\xi' \leqq \xi \leqq \xi'', \qquad 0 \leqq \eta < 2\pi,$$

whose hyperbolic area is

$$\int_0^{2\pi} \int_{\xi'}^{\xi''} \frac{d\xi\, d\eta}{\xi^2} = 2\pi\left(\frac{1}{\xi'} - \frac{1}{\xi''}\right) < \frac{2\pi}{\xi'} < \frac{2\pi}{K^2}.$$

Thus the spherical area is less than 2π, and this gives a contradiction.

PICARD'S SECOND THEOREM. *If a function is regular and single-valued near an essential singular point, it assumes any finite value, with at most one exception, in an arbitrarily small neighborhood of that point.*

In that statement, we have transformed the region $1 < |z| < \infty$ into a small neighborhood of the given essential singular point, with the exception of that point itself. Similarly, we can obtain from Theorem 5, for each of the five cases in §2, a statement about the values assumed by a meromorphic function near an essential singularity. For example, the first case there gives the following result. *Let $w_1, \cdots, w_5$ be any five distinct complex numbers. Then if $f(z)$ is meromorphic and single-valued near an essential singular point, at least one of the five equations $f(z) = w_k$ will have a simple root in an arbitrarily small neighborhood of that point.*

8. **Definitiveness of the results.** We now raise the question whether any of our theorems can be replaced by better theorems of the same type; i.e., theorems in which the same conclusions are obtained from a weaker hypothesis about the orders with which the exceptional values are assumed. We shall show that this is not the case.

In doing this, we need to make use of certain well known special functions, which however cannot be considered elementary. Consequently, this section forms an exception to the statement at the beginning of the paper concerning the material assumed as known.

Consider first Theorem 2, which is a generalization of Picard's first theorem. We need to show that for each possible set of exceptional values, of preassigned orders, such that the total weight is $\leqq 2$, a meromorphic function can be found which assumes the exceptional values only with exactly those orders, and which is not a constant. If there are only one or two exceptional values, at ∞ or at 0 and ∞, then the function e^z may be used, since it does not assume either of those values. Consider now the case in which there are three exceptional values. If two of these are of the second order, at 1 and -1, and the third of any order, at ∞, then we may use the function $\cos z$. The only other cases to be considered with three exceptional values are those in which the orders are (2, 3, 3), (2, 3, 4), (2, 3, 5), (2, 3, 6), (2, 4, 4), and (3, 3, 3). In these cases we may use Schwarz's triangle functions, which in the first three cases are rational functions, and in the last three cases are elliptic functions. Finally, if there are more than three exceptional values, we can have only the case (2, 2, 2, 2). Here the function $\wp(z)$ may be used, since it has poles of the second order at the period points, and assumes the values e_1, e_2, e_3 at the half-period points, with the multiplicity 2; and the values ∞, e_1, e_2, e_3 are not assumed at any other points.

Thus we can find a meromorphic function which assumes the exceptional values with the desired multiplicities. But we should also show that the positions of the exceptional values are arbitrary. If there are not more than three exceptional values, they can be taken into any positions by a linear trans-

formation, in the plane of the dependent variable. The only case which remains is the case (2, 2, 2, 2). Now it is shown in the theory of elliptic modular functions that, by a proper choice of periods, the quotient $(e_2 - e_3)/(e_1 - e_3)$ may be made to assume any complex value except 0, 1, and ∞. From this, it is easily seen that by a proper choice of periods and of complex numbers a, b, c, d, we may determine a function $(a\wp(z) + b)/(c\wp(z) + d)$, which has any four given distinct complex numbers as exceptional values.

Hence, in the case of Theorem 2, no improvement of the kind mentioned is possible. It is then evident that the same is true for Theorem 1, since Theorem 2 was deduced from it; it is of course immaterial what circle C is. Likewise, we shall know that Theorem 3 cannot be improved, if we can deduce Theorem 1 (with C the unit circle) from it.

We assume then the truth of Theorem 3, and try to prove from it that $d(w) \leqq K(\delta)d[z]$. It will be sufficient to show this for the case $z = 0$ and $w = 0$. For otherwise, the point z under consideration can be taken to the origin by a linear transformation which leaves the unit circle fixed, so that $d[z]$ is invariant; and the point w being considered can be taken to 0 by a rotation of the sphere, which leaves $d(w)$ invariant. We then have $w = f(z)$ with $f(0) = 0$; and we have only to prove that $|f'(0)|$ is less than a bound which depends only on δ. Now since all such functions, with a given δ, are supposed to form a normal family, they are equicontinuous in the interior of the unit circle. Hence we can find δ', which depends only on δ, so that $|f(z)| < 1$ for $|z| < \delta'$. From this, we conclude that $|f'(0)| \leqq 1/\delta'$, and this completes the proof.

In the case of Theorem 4, we are concerned with regular functions. The cases to be considered here are those in which the finite exceptional values have total weight $\leqq 1$; that is, there is one finite exceptional value, of any order, or there are two finite exceptional values, each of the second order. In these cases, we make use of the functions $f(z) = ae^{nz} + b$ and $f(z) = a \cos(nz + c) + b$. The constants a and b, or a, b, and c, with $a \neq 0$, can be chosen so that the finite exceptional values and $f(0)$ are arbitrary. For $n \to \infty$, neither of these functions remains bounded in any circle about the origin; i.e., Theorem 4 cannot be extended to these cases.

Finally, in the case of Theorem 5, we need to find functions having essential singularities at ∞ and satisfying the same conditions with regard to exceptional values as in the case of Theorem 2, at the beginning of this section. Now the function $f(z)$ constructed there can also be used here, except in the cases in which it is a rational function; then it would not have the required essential singular point, but we could use $f(e^z)$ in its place.

Thus, in all cases, our theorems are the definitive results of their kind.

UNIVERSITY OF CALIFORNIA.

CERTAIN EQUIVALENCES AND SUBSETS OF A PLANE

By F. Burton Jones

In the following paragraph is stated a sequence of propositions which are known to hold true in the plane and in certain other spaces.[1] It is the purpose of this paper to study the relations that exist between these propositions in spaces which satisfy R. L. Moore's Axioms 0–4.[2] For example, it will be shown that in any space satisfying these axioms, each proposition is a consequence of each succeeding proposition and that those propositions of like number are equivalent. Furthermore, in any space which satisfies Moore's Axioms 0–4 and the axiom stated in the second section of this paper, all of these propositions are equivalent. A space exists, however, as is shown by an example, in which all of these propositions are false but which satisfies Axioms 0–4 and the author's axiom of §2 and which is a subset of a plane.

Let α denote the following sequence of propositions:

Proposition 1. *If A and B are two distinct points and neither of the two mutually exclusive closed and compact point sets H and K separates A from B, then $H + K$ does not separate A from B.*

Proposition 2a. *If A and B are distinct points, there exists a simple closed curve separating A from B.*

Proposition 2b. *If the common part of the closed and compact point sets H and K is a continuum and neither H nor K separates the point A from the point B, then $H + K$ does not separate A from B.*

Proposition 2c. *No arc separates space.*

Proposition 3a. *If A and B are distinct points, $A + B$ is a subset of a simple closed curve.*

Proposition 3b. *Space contains no end points.*

Proposition 3c. *If P is a point, there exists a simple closed curve containing P.*

Proposition 4a. *Space contains no local cut points.*

Proposition 4b. *If P is a point of a connected domain, then $D - P$ is connected.*

Proposition 5. *If A is a point of a region R and B is a point distinct from A, there exists in R a continuum T separating A from B.*

Proposition 6a. *If A is a point of a region R and B is a point distinct from A, then R contains a compact continuum separating A from B.*

Received September 16, 1938; presented in part to the American Mathematical Society, June 18, 1936.

[1] See Chapter IV of R. L. Moore's *Foundations of Point Set Theory*. This book will hereinafter be referred to as *Foundations*. For the definition of certain terms and phrases used without definition in this paper the reader is referred to *Foundations*.

[2] Loc. cit.

PROPOSITION 6b. *If A is a point of a region R and B is a point distinct from A, then R contains a simple closed curve separating A from B.*[3]

1. **Consequences of Axioms 0–4.** Let S denote the set of all points.[4]

THEOREM 1. *Proposition* 2a *implies Proposition* 1.

Proof. Suppose, on the contrary, that H and K are two mutually exclusive closed point sets neither of which separates the point A from the point B such that $H + K$ does separate A from B. It follows from Proposition 2a and the arguments in *Foundations* for Theorems 9, 21, 22, 23, and 24 of Chapter IV that $H + K$ contains a continuum T separating A from B. Since H and K are mutually separate, T is a subset of one of them. Consequently, this set separates A from B contrary to hypothesis.

However, *Proposition* 1 *does not imply Proposition* 2a as is shown by the following example.

Example 1. Let S denote a subset of the number plane consisting of the points $(0, 1)$ and $(0, -1)$ together with all points whose abscissas are positive except those points (X, Y) such that $X = n^{-1}$ $(n = 1, 2, 3, \cdots)$ and $1\frac{1}{2} \geqq Y > 1$ or $1 > Y \geqq \frac{1}{2}$ or $-\frac{1}{2} \geqq Y > -1$ or $-1 > Y \geqq -1\frac{1}{2}$.[5] It is apparent with the help of Theorem 9 on page 96 of *Foundations* that S satisfies Axioms 0–4. Furthermore, Proposition 1 holds true in S. But Proposition 2a does not hold true, for there exists no simple closed curve separating $(0, 1)$ from $(0, -1)$.

THEOREM 2. *Proposition* 2a *implies Proposition* 2b.

Proof. Suppose, on the contrary, that $H + K$ separates A from B. Let AXB and AYB denote arcs from A to B lying in $S - K$ and $S - H$, respectively, and let U denote the component of $S - (H + K)$ which contains A. Let A' denote the last point of AXB belonging to $AXB \cdot AYB \cdot U$ and let B' denote the first point of the interval $A'B$ of AXB which is in AYB. It is clear that B' does not belong to $U + H + K$, and that $H + K$ separates A' from B'. Denote the interval $A'B'$ of AXB by $A'OB'$, where O is a point of H, and denote the interval $A'B'$ of AYB by $A'PB'$, where P is a point of K. These intervals have only their end points in common. Let J denote the simple closed curve $A'OB' + A'PB'$. The simple closed curve J does not intersect $H \cdot K$. Let D denote the complementary domain of J which does not contain $H \cdot K$, and let L denote an interval of $A'PB'$ which contains $A'PB' \cdot K$ but which contains neither A' nor B'. It follows from Proposition 2a and the argument in *Foundations* for Theorem 9 of Chapter IV, that there exists a simple closed curve C separating O from P and containing no point of $K + L + (D + J) \cdot H + A'OB'$. By part

[3] Axiom 5 of *Foundations*.

[4] If the reader is not familiar with *Foundations*, it may facilitate his reading of this paper to know that *every connected open subset of S is arc-wise connected.* Furthermore, *any locally connected complete metric space which contains no cut points and in which the Jordan Curve Theorem holds true satisfies Axioms* 0–4.

[5] These sets are removed so that there will exist no simple closed curve through $(0, 1)$ and $(0, -1)$ which would separate S into more than two domains contrary to Axiom 4. The points $(n^{-1}, \pm 1)$ are retained so that S will be connected im kleinen at $(0, 1)$ and $(0, -1)$.

(3) of the conclusion of Theorem 11 on page 163 of *Foundations* P is separated from O, on J, by two points which are the extremities of a segment T of C which lies wholly in D. Obviously, T contains no point of $H + K$. Hence $T + A'PB'$ contains an arc from A' to B' not intersecting $H + K$, and this is a contradiction.

THEOREM 3. *Proposition* 2b *implies Proposition* 2c.

The argument for Theorem 19 on page 191 of *Foundations* without change establishes Theorem 3.

THEOREM 4. *Proposition* 2c *implies Proposition* 2a.

Proof. Let AXB denote an arc from A to B. There exists in $S - X$ an arc from A to B which contains an interval $A'YB'$ which has only the points A' and B' in common with AX and XB, of AXB, respectively. If the simple closed curve $A'XB'$ (of AXB) $+ A'YB'$ does not contain A nor separate A from B, then let WZ denote an arc such that (1) $WZ - Z$ lies in the complementary domain of $A'XB' + A'YB'$ which contains A, and (2) $WZ \cdot AXB = Z$, an internal point of $A'XB'$. By Proposition 3c, $S - AXB$ is connected. Hence there exists in $S - AXB$ an arc WY from W to Y. Then $WY + WZ$ contains a segment T which lies in the complementary domain of $A'XB' + A'YB'$ containing A such that one end point of T is Z and the other end point is a point of $A'YB'$. It follows with the help of Theorems 3 and 4 on page 155 of *Foundations* that $T + A'XB' + A'YB'$ contains a simple closed curve separating A from B. A similar argument may be given in case $A'XB' + A'YB'$ contains A or $A + B$.[6]

THEOREM 5. *Propositions* 2a, 2b, *and* 2c *are equivalent.*

Theorem 5 follows from Theorems 2, 3, and 4.

THEOREM 6. *Proposition* 3a *implies Proposition* 2a.[7]

However, *Proposition* 2a *does not imply Proposition* 3a as is shown by Example 2.

Example 2. Let S denote the subset of Example 1 consisting of those points of Example 1 whose ordinates are positive. Axioms 0–4 are satisfied. It is easy to see that Proposition 2a holds true in S, but Proposition 3a does *not* hold true, since S contains no simple closed curve containing (0, 1).

THEOREM 7. *Proposition* 3a *implies Proposition* 3b.

THEOREM 8. *Proposition* 3b *implies Proposition* 3c.

Proof. By Proposition 3b there exists an arc APB containing P as an internal point. There exists an arc AXB from A to B lying in $S - P$. Let A' denote the last point of AXB in the interval AP of APB, and let B' denote the first point of AXB which follows A' and is in APB. The sum of the intervals AXB and APB from A' to B' is a simple closed curve containing P.

[6] This argument shows that *if A and B are two distinct points and there exists an arc from A to B which does not separate space, then there exists a simple closed curve separating A from B.*

[7] If a proof of Theorem 6 does not occur to the reader, he is referred to the proof of Theorem 2 on page 57 of my paper *Concerning certain topologically flat spaces*, Trans. Amer. Math. Soc., vol. 42(1937), pp. 53-93. As a matter of fact, this argument shows that *any two points of a simple closed curve may be separated by a simple closed curve.*

THEOREM 9. *Proposition* 3c *implies Proposition* 3a.[8]

Proof. Since A lies on some simple closed curve, it is not a local end point.[9] Likewise, B is not a local end point. By Axiom 3, neither A nor B is a cut point and no point separates A from B. Hence, by Theorem A′ of the footnote on page 56 of the author's paper *Concerning certain topologically flat spaces*[10] there exists a simple closed curve containing $A + B$.

THEOREM 10. *Propositions* 3a, 3b, *and* 3c *are equivalent.*

Theorem 10 follows from the three preceding theorems.

THEOREM 11. *Proposition* 4a *implies Proposition* 3a.

Theorem 11 may be established by a slight modification of the proof of Theorem 9.

A simple example shows that *Proposition* 3a *does not imply Proposition* 4a.

Example 3. Let S denote the subset of the number plane obtained by removing the points of the interval of the X-axis from -1 to $+1$ except the origin. Obviously, Axioms 0–4 are satisfied and Proposition 3a holds true in S. However, the origin is a local cut point of S.

THEOREM 12. *Propositions* 4a *and* 4b *are equivalent.*

Proof. That Proposition 4b implies Proposition 4a is evident. It remains only to show that Proposition 4a implies Proposition 4b. Suppose, on the contrary, that there exists a connected domain Q containing a point P such that $Q - P$ is the sum of two mutually separate sets H and K. Both H and K have P as a limit point. Let R denote a region containing P and lying in Q. Hence, if D is any connected domain whatever lying in R and containing P, $D - P = D \cdot H + D \cdot K$ is disconnected. This contradicts Proposition 4a.

THEOREM 13. *Proposition* 5 *implies Proposition* 4a.

Proof. Suppose that S contains a local cut point P. Then there exists a connected domain D such that $D - P$ contains two components. Let X and Y denote points belonging to different components of $D - P$, and let N denote an arc from X to Y lying in $S - P$. There exists in $D - D \cdot N$ a region R which contains P and by Proposition 5, $R - P$ contains a continuum T separating P from X. Since T contains no point of N, T also separates P from Y. This is, however, impossible since X and Y are in different components of $D - P$ and D is connected.

It may appear at first sight that Proposition 4a implies Proposition 5 on this basis. Such is *not* the case as the following example shows.

Example 4. Let S denote a subset of the number plane obtained by removing all of the points (X, Y) such that $X = 0$ or $\pm n^{-1}$ $(n = 1, 2, 3, \cdots)$ and $0 < Y \leqq 1$. Again, S satisfies Axioms 0–4. Furthermore, S contains no local cut

[8] The argument for this theorem establishes a somewhat stronger result: *If A and B are two distinct non-end points of S, then there exists a simple closed curve containing $A + B$.*

[9] As defined by G. T. Whyburn, a point X is a local end point provided that there exist a domain V containing X and a sequence $V_1, V_2, V_3, \cdots$ of domains such that $\Pi V_i = X$ and for each i, the boundary of V_i with respect to V is a single point. See the reference of footnote 19.

[10] Loc. cit.

points. But if R is a circular region of diameter one, with center at (0, 0), $R - (0, 0)$ does not contain a continuum T separating (0, 0) from (2, 0). Hence, *Proposition* 4a *does not imply Proposition* 5.

THEOREM 14. *Proposition* 6a *implies Proposition* 5.

The converse, however, is not true.

Example 5. Let S denote the subset of the number plane obtained by removing the points of the interval from 0 to 1 of the Y-axis except the origin. Axioms 0–4 and Proposition 5 hold true in S but Proposition 6a fails to hold at (0, 0). Consequently, *Proposition* 5 *does not imply Proposition* 6a.

THEOREM 15. *Propositions* 6a *and* 6b *are equivalent.*

Proof. It is evident that Proposition 6b implies Proposition 6a. That Proposition 6a implies 6b has been shown by Theorem 13 of a paper by the author entitled *Concerning R. L. Moore's Axiom* 5.[11]

THEOREM 16. *Any proposition of the sequence* α *implies each proposition of the sequence which precedes it.*

There are many theorems, especially those of Chapter IV of *Foundations*, whose relations to the propositions of α are of considerable interest. Of particular interest are the following three propositions.

PROPOSITION 1′. *Suppose that* (1) J *is a simple closed curve,* (2) XY *is an arc such that* $J \cdot XY$ *is* X, *and* (3) I *is the complementary domain of* J *which contains* Y. *Then* $I - I \cdot XY$ *is connected.*

PROPOSITION 4′a. *Every point of the boundary of a complementary domain of a compact continuous curve is accessible from that domain.*[12]

PROPOSITION 4′b. *Suppose that* H *and* K *are two compact continua such that* (a) K *does not separate any two points of* H *from each other,* (b) T, *the common part of* H *and* K, *is totally disconnected, and* (c) $K - T$ *is connected. Then there exists a simple closed curve which separates* $H - T$ *from* $K - T$ *and which contains* T *but no point of* $(H + K) - T$.[13]

THEOREM 17. *Proposition* 2a *implies Proposition* 1′.

Proof. Suppose the contrary. Then $I - XY$ contains two components containing points W and Z, respectively. By Proposition 2c, XY does not separate W from Z. The set $J \cdot XY$ is connected and J does not separate W from Z. Hence, by Proposition 2b, $J + XY$ does not separate W from Z, and this is a contradiction.

On the other hand, Example 1 shows that *Proposition* 1′ *does not imply Proposition* 2a. That *neither Propositions* 1 *nor* 1′ *implies the other* is shown by Examples 6 and 7.

Example 6. Let S denote the subset of the number plane consisting of (0, 0) together with all of the points of Example 2. It is easy to see that Axioms 0–4 and Proposition 1 hold true in S. However, if J is a simple closed curve containing (0, 0) and XY is an arc irreducible from J to (0, 1), then $XY - X$

[11] Bull. Amer. Math. Soc., vol. 44(1938), pp. 689–692.

[12] See Theorem 56 on page 229 of *Foundations*.

[13] See Theorem 42 on page 217 of *Foundations*.

separates the complementary domain of J which contains Y. Hence *Proposition* 1 *does not imply Proposition* 1′.

Example 7. Except for the four points $(\pm 2, \pm 1)$, remove from the number plane the points of the two squares whose vertices are (2, 2), (2, −2), (6, 2), (6, −2) and (−2, 2), (−2, −2), (−6, 2), (−6, −2), and remove the points of the interiors of these squares. Then, except for the points on the lines $Y = \pm 1$, remove all points (X, Y) such that $X = \pm(2 - n^{-1})$ $(n = 1, 2, 3, \cdots)$ and $\frac{1}{2} \leqq \pm Y \leqq 1\frac{1}{2}$. Let S denote the subset of the plane remaining. Axioms 0–4 hold true in S. Let H and K denote the intervals of the lines $Y = 1$ and $Y = -1$ from $X = -2$ to $X = 2$. Neither H nor K separates S, but $H + K$ does. Hence, Proposition 1 does not hold true in S. However, Proposition 1′ holds true in S because no simple closed curve contains (2, 1), (2, −1), (−2, 1) or (−2, −1). Consequently, *Proposition* 1′ *does not imply Proposition* 1.

LEMMA. *Proposition* 4′a *implies Proposition* 2a.

Proof. Suppose, on the contrary, that Proposition 4′a holds true but Proposition 2a does not. Then there exist two distinct points A and B such that no simple closed curve separates A from B. Let AXB denote an arc from A to B. By the argument for Theorem 4, AXB separates S. Let D denote one of the complementary domains of AXB, and let A' and B' denote the first and last points, respectively, that AXB has in common with the boundary of D. Then the interval $A'B'$ of AXB from A' to B' separates space. It follows from Proposition 4′a and Theorem 1 on page 86 of *Foundations* that there exists in D an arc segment T whose end points are A' and B'. Let J denote $T + A'B'$. It is clear from Axiom 4 that $A'B'$, being a subset of the simple closed curve J, does not separate space, and this is a contradiction.

THEOREM 18. *Proposition* 4′a *implies Proposition* 3a.

Proof. Let AXB denote an arc from A to B. By the preceding lemma, $S - AXB$ is connected. Since both A and B are accessible from $S - AXB$, there exists in $S - AXB$ an arc segment T whose end points are A and B. Then $T + AXB$ is a simple closed curve containing $A + B$.

THEOREM 19. *Proposition* 4′b *implies Proposition* 3a.

Proof. Let P denote any point of S and let Q denote a point distinct from P. There exists an arc PXQ from P to Q. Let $H = P$ and $K = PXQ$. Then by Proposition 4′b, there exists a simple closed curve separating $H - P$ from $K - P$ and which contains P but no point of $(H + K) - P$. Hence, Proposition 3c holds true. By Theorem 10, Proposition 3a holds true.

That *neither Proposition* 4′a *nor Proposition* 4′b *implies Proposition* 4a is shown by Example 3.

That *Proposition* 4a *implies neither Proposition* 4′a *nor Proposition* 4′b is shown by Example 4. In Example 4, let J denote the rectangle whose vertices are (2, 0), (−2, 0), (2, 2) and (−2, 2). Then (0, 0) is not accessible from the interior of J. Hence, Proposition 4′a does not hold true. Likewise, if H and K denote the two arcs of J from (0, 0) to (0,2), there exists no simple closed curve separating $H - H \cdot K$ from $K - H \cdot K$. Hence, Proposition 4′b does not hold true.

THEOREM 20. *Proposition* 6a *implies Propositions* 4′a *and* 4′b.[14]

Theorem 20 follows from Theorems 42 and 56 on pages 217 and 229, respectively, of *Foundations* and Theorem 15.

2. **Equivalences on the basis of Axioms 0–4 and $\bar{5}$.**

AXIOM $\bar{5}$. *If A is a point of a region R and B is a point distinct from A, there exists in R a closed and compact point set T separating A from B.*

THEOREM 21. *Proposition* 1 *implies Proposition* 6a.

Proof. By Axiom $\bar{5}$, R contains a closed and compact point set T separating A from B. Let D denote the component of $S - T$ which contains A and let Q denote the component of $S - \bar{D}$ which contains B.[15] The boundary β of Q is closed and compact, and separates A from B. Furthermore, β is not the sum of two mutually exclusive closed sets; for if it were, neither of them would separate A from B and, by Proposition 1, their sum β would not separate A from B. But if β is not the sum of two mutually exclusive closed sets, it is, by Theorem 30 on page 18 of *Foundations*, connected. Hence β is a compact continuum lying in R and separating A from B.

THEOREM 22. *Propositions* 4′a *and* 4′b *and all of the propositions of the sequence α are equivalent.*

Theorem 22 is an immediate consequence of Theorems 16, 18, 19, 20, and 21.

If the reader will notice that in all of the examples given so far, in which some of the propositions of α are false, Axiom $\bar{5}$ does not hold true; then he will not find it surprising that the author at one time thought that all of these propositions were equivalent on this basis because they all held true. Such is *not* the case *even though space be a subset of a plane*. There follows an example of a subset of a plane satisfying Axioms 0–4 and $\bar{5}$ in which all of these propositions are false.

Definition. Suppose that M is a subset of a plane E. Select in E a line L. A set is said to be a *horizontal component* of $E - M$ provided that it is a maximal connected subset of the common part of $E - M$ with L or a line parallel to L. If I is a non-degenerate horizontal component of $E - M$, I is obviously either an interval or an interval less one or both of its end points.

Definition of processes P and Q and the sets $P \circ M$ and $Q \circ M$. Suppose that M is a subset of a plane E, and suppose that I lies in $E - M$ and is either an interval or an interval less one or both of its end points. In order to facilitate this description, set up a rectangular coördinate system in E with (0, 0) at one end of I and (1, 0) at the other. Following the usual method of constructing the Cantor Set, let S_{01} denote the middle $\frac{1}{3}$-segment of I; let I_{11} and I_{12} denote the left and right components of $I - S_{01}$, respectively; let S_{11} and S_{12} denote the middle $\frac{1}{3}$-segments of I_{11} and I_{12}, respectively; let I_{21} and I_{22} denote the left and right components of $I_{11} - S_{11}$, respectively; and let I_{23} and I_{24} denote the left

[14] The relation between Propositions 4′a, 4′b and 5 will not be discussed in this paper. I have discovered examples to show that no one of these propositions implies either of the other two.

[15] The notation $\bar{D}$ means D plus its limit points.

and right components of $I_{12} - S_{12}$, respectively; etc., indefinitely. The reader should note that, for a particular value of m, the maximum value of n for which I_{mn} is defined is 2^m, and that the length of I_{mn} is 3^{-m}. For each m and n ($m = 1, 2, 3, \cdots$; $n = 1, 2, \cdots, 2^m$) let X_{mn} denote the abscissa of the mid-point of I_{mn}.

By *applying process P to M at I* shall be meant removing from M all points (X, Y) of M such that either (1) $(X, 0)$ is a point of I which is an end point of some I_{mn} and $-\frac{1}{6} \leqq Y \leqq \frac{1}{6}$ or (2) $(X, 0)$ is internal to infinitely many of the sets I_{mn} and $-\frac{1}{6} < Y < \frac{1}{6}$. (See Figure 1a.) Let $P \circ M \circ I$ denote the subset of M remaining after applying P to M at I, and let $P \circ M$ denote $\Pi P \circ M \circ I$ for all non-degenerate horizontal components I of $E - M$.

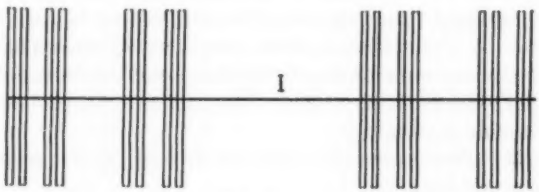

FIG. 1a

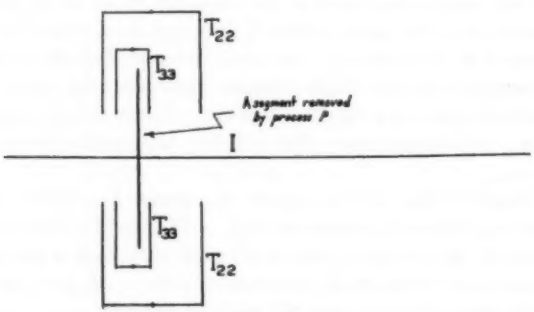

FIG. 1b

For each m and n ($m = 2, 3, 4, \cdots$; $n = 2, 3, \cdots, 2^m - 1$) let T_{mn} denote the set of all points (X, Y) of M such that (1) $X = X_{mn} \pm \frac{5}{6}\, 3^{-m}$ and $\frac{1}{12} \leqq \pm Y \leqq \frac{1}{6} + 3^{-m}$ or (2) $|X - X_{mn}| \leqq \frac{5}{6}\, 3^{-m}$ but $X \neq X_{mn}$ and $Y = \pm[\frac{1}{6} + 3^{-m}]$. (See Figure 1b.) By *applying process Q to M at I* shall be meant removing from M the sum of all of the sets T_{mn}. Denote the subset of M remaining by $Q \circ M \circ I$ and let $Q \circ M$ denote $\Pi Q \circ M \circ I$ for all non-degenerate horizontal components I of $E - M$.

Example 8. Let M denote the set of all points of the number plane E except points of the interval $(-1, 0)$ to $(1, 0)$ of the X-axis distinct from the origin, and let L denote the X-axis. Let $M_1 = (P \circ M) \cdot (Q \circ M)$, $M_2 = (P \circ M_1) \cdot (Q \circ M_1)$, $M_3 = (P \circ M_2) \cdot (Q \circ M_2)$, etc., indefinitely. Let $M_\infty = M_1 \cdot M_2 \cdot M_3 \cdot \cdots$. The

reader should have no difficulty in seeing that M is an inner-limiting set and that applying processes P and Q to M at each of the horizontal components of $E - M$ removes from M a set which is the sum of a countable number of closed sets. In fact, for each i, $E - M_i$ is the sum of a countable number of closed sets and M_i is an inner-limiting set. Hence, M_∞ is an inner-limiting subset of the plane and Axioms 0 and 1 hold true in M_∞ (if M_∞ is considered as a space)[16] leaving "limit point" invariant. However, M_∞ is not connected im kleinen at all of its points; for example, the points of the Y-axis between $(0, -\frac{1}{6})$ and $(0, \frac{1}{6})$. Let H denote the set of all points of M_∞ at which M_∞ is not connected im kleinen. Every component of H is a segment of either the Y-axis or a line parallel to the Y-axis. Let G_1 denote the collection of all intervals AB such that $AB - (A + B)$ is a component of H. Let G_2 denote the collection of all point sets g such that the only point in g is a point of $E - G_1^*$.[17] Then $G = G_1 + G_2$ is an upper semi-continuous collection of compact continua filling up E, and no element of G separates E. If "point" is interpreted to mean an element of G and "region" is interpreted in a natural way, then G is a plane.[18] Let S denote the subset of G corresponding to M_∞ of E. It is clear that, considered as a space with "limit point" as in G, S satisfies Axioms 0–3. Furthermore, if one observes that no simple closed curve lying in S contains more than one boundary point of a component of $G - S$, it is clear that Axiom 4 holds true in S. Axiom $\bar{5}$ is satisfied, for

(1) if A is an element of G_1, then the "set T" of Axiom $\bar{5}$ can be the sum of two arcs;

(2) if A is a "point" on the boundary of a component of $G - S$ but not in G_1, then the "set T" can be an arc; and

(3) if A is a "point" of G_2 in S which is not on the boundary of a component of $G - S$, then the "set T" can be a simple closed curve. But obviously S contains a local cut point and consequently, by Theorem 22, *none of the propositions of α holds true in S.*

The reader will notice that in Example 8 Proposition 1′ does not hold true. However, *Proposition* 1′ *does not imply Proposition* 6a even though S is a subset of a plane. The following is an indication of an example of a subset of a plane satisfying Axioms 0–4 and $\bar{5}$ in which Proposition 1′ holds true but in which none of the propositions of α holds true.

Indication of Example 9. (1) As in the construction of Example 8, remove from the number plane E all of the points of E of the interval from $(-1, 0)$ to $(1, 0)$ of the X-axis except the origin, and let M denote the set remaining. (2) Remove from M a simple infinity of simple domains plus their boundaries, $J_1, J_2, J_3, \cdots$, as indicated by the shaded portions of Figure 2, and denote the subset of M remaining by M_1. I. Apply processes P and Q to M_1 at each of the

[16] Theorem 120 on page 83 of *Foundations*.

[17] G_1^* denotes the sum of the elements of G_1.

[18] See the discussion on page 326 of *Foundations*. Also see Theorems 2 and 6 on pages 369 and 434, respectively, of *Foundations*.

intervals $I_1, I_2, I_3, \cdots$ of Figure 2, and denote the subset of M_1 remaining by M_{11}. II. Somewhat as in (2), at every one of the sets T_{mn} produced by applying

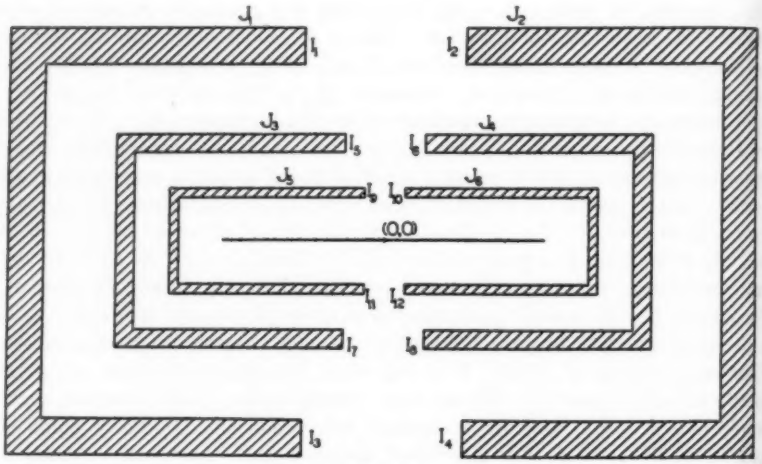

FIG. 2

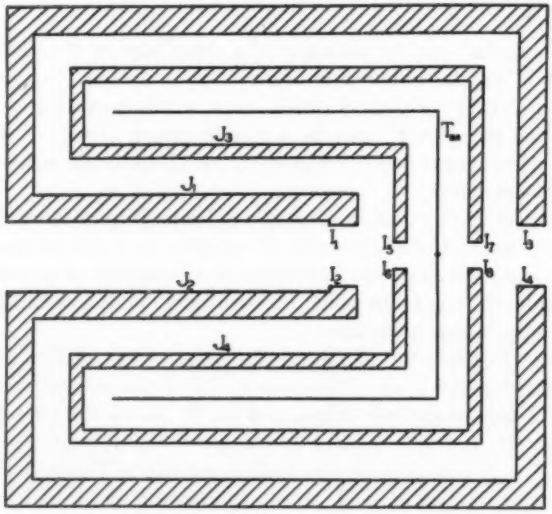

FIG. 3

process Q in I, remove from M_{11} a simple infinity of simple domains plus their boundaries, $J_1, J_2, J_3, \cdots$, as indicated by Figure 3, and denote the subset of

M_{II} remaining by M_2. By repeating I and II on M_2, a subset M_3 of M_2 is obtained. In this way an infinite sequence $M_1, M_2, M_3, \cdots$ of inner-limiting subsets of E is constructed whose common part S is also an inner-limiting subset of E. *The subset S of E is a space satisfying Axioms* 0–4 *and* $\bar{5}$ *in which none of the propositions of α holds true.* But if C is a component of $E - S$, then either $\bar{C} \cdot S$ contains only one point, a local cut point of S, or every point of $\bar{C} \cdot S$ is an end point of S. Hence, *Proposition* 1′ *holds true in S.*

3. **Certain locally divisible spaces.**

Definition. A space S is *locally remotely connected* provided that, if P is a point of a region R, there exists in R a domain D containing P such that $S - \bar{D}$ is connected.

Definition. A space S is *locally divisible* provided that, if P is a point of a region R, there exists in R a domain D containing P such that $S - \bar{D}$ is the sum of a finite number of connected sets.[19]

THEOREM 23. *In order that a locally divisible space which either is locally connected and metric or satisfies Moore's Axioms* 0–2 *be locally peripherally compact, it is necessary and sufficient that it satisfy Axiom* $\bar{5}$.

Proof. It is evident that the condition is necessary and it remains only to prove that the condition is sufficient. Suppose that P is a point of a region R. Then there exists in R a domain D such that the complement of $\bar{D}$ is the sum of a finite number of connected sets $C_1, C_2, \cdots, C_n$. For each i $(1 \leqq i \leqq n)$, there exists in D a closed and compact point set T_i which separates P from a point of C_i. The complementary domain of $\sum T_i$ which contains P lies in D and has a compact boundary. Hence the space is locally peripherally compact.

THEOREM 24. *Any connected, locally peripherally compact space which either is locally connected and metric or satisfies Moore's Axioms* 0–2 *is locally divisible.*

Proof. If P is a point of a region R, there exists in R a domain V whose boundary β is compact. Let U denote a region containing P such that $\overline{U}$ is a subset of V, and let G denote the collection of all components of the complement of $\overline{U}$. The elements of G are mutually exclusive domains and G covers β. By Theorem 22 on page 15 of *Foundations*, G contains a finite subcollection H covering β. Let D denote the complement of $\overline{H^*}$. Then D is a domain in R containing P such that the complement of $\bar{D}$ is the sum of a finite number of connected sets.

[19] G. T. Whyburn, *The cyclic and higher connectivity of locally connected spaces*, Amer. Jour. Math., vol. 53(1931), pp. 427–442.

If S is a space satisfying Moore's Axioms 0–1, we have the following two theorems, the proofs of which will be left to the reader.

THEOREM. *In order that S be locally remotely connected it is necessary and sufficient that if P is a point of a region R, there exists in R a domain D containing P such that $S-D$ is connected.*

THEOREM. *In order that S be locally divisible it is necessary and sufficient that if P is a point of a region R, there exists in R a domain D containing P such that $S-D$ is the sum of a finite number of connected sets.*

THEOREM 25. *In order that a connected space satisfying Axiom* $\bar{5}$ *which either is locally connected and metric or satisfies Moore's Axioms* 0-2 *be locally divisible, it is necessary and sufficient that it be locally peripherally compact.*

Theorem 25 is an immediate consequence of Theorems 23 and 24.

THEOREM 26. *Any connected, locally connected, locally divisible metric space which satisfies Axiom* $\bar{5}$ *is completely separable.*

Theorem 26 follows from the preceding theorem and a theorem of the author's found elsewhere.[20]

THEOREM 27. *Any locally divisible space which satisfies Moore's Axioms* 0–3 *is locally remotely connected.*

Proof. If P is a point of a region R, there exists in R a domain D such that the complement of $\bar{D}$ is the sum of a finite number of connected sets $C_1, C_2, \cdots, C_n$. For each i $(1 \leqq i \leqq n)$, let N_i denote an arc containing a point of C_1 and a point of C_i but not containing P. The complement of $\sum_1^n (\bar{C}_i + N_i)$ is a domain which lies in R and contains P and whose complement is connected. It follows from a theorem stated in footnote 19 that the space is locally remotely connected.

THEOREM 28. *Any locally divisible space which satisfies Axiom* $\bar{5}$ *and Moore's Axioms* 0–3 *is cyclically connected.*

This follows from Theorem A of the author's paper on topologically flat spaces[21] and Theorem 25.

THEOREM 29. *All of the propositions of* α *hold true in any locally divisible space satisfying Axiom* $\bar{5}$ *and Moore's Axioms* 0–4; *furthermore, such a space is locally simple, i.e., if* P *is a point of a region* R *there exists in* R *a simple domain containing* P.

Proof. The first part of the theorem follows from Theorems 22 and 28. That the space is locally simple may be established by the following argument. By Theorem 27, if P is a point of a region R, there exists in R a domain D whose complement C is connected. By Proposition 6b, there exists in the complement of $\bar{C}$ a simple closed curve J separating P from a point of C. The complementary domain of J which contains P lies in R.

THEOREM 30. *In order that a metric space* S *satisfying Axiom* $\bar{5}$ *and Moore's Axioms* 0–4 *be a subset of a plane or a sphere in which the propositions of* α *hold true, it is necessary and sufficient that* S *be locally remotely connected.*

THEOREM 31. *In order that a metric space* S *satisfying Axiom* $\bar{5}$ *and Moore's Axioms* 0–4 *be a subset of a plane or a sphere in which the propositions of* α *hold true, it is necessary and sufficient that* S *be locally divisible.*

Proof of Theorems 30 *and* 31. The conditions are necessary. For simplicity suppose that S is a subset of a plane and that all of the propositions of α hold true in S. Then if A is a point of a region R of S, A is the center of a circle C of the plane such that if I denotes the interior of C, then $(C + I) \cdot S$ is a subset

[20] *A connected, locally connected, locally peripherally separable metric space is completely separable.* See my paper, *A theorem concerning locally peripherally separable spaces*, Bull. Amer. Math. Soc., vol. 41(1935), pp. 437–439.

[21] See footnote 7.

of R. Let B denote a point of $C \cdot S$. By Proposition 6b, there exists a simple closed curve J lying in $I \cdot S$ and separating A from B. Let D denote the component of $S - J$ which contains A. The domain D does not intersect C and is, therefore, a subset of $I \cdot S$. Hence, D lies in R. It follows from Axiom 4 that $S - \bar{D}$ is connected.

That the conditions are sufficient follows from Theorems 26 and 29 and a theorem of J. H. Roberts'.[22]

THEOREM 32. *Any locally divisible space satisfying Moore's Axioms* 0–3 *in which Proposition* 5 *holds true satisfies the author's Axiom* 5_1^*.[23]

Proof. Suppose that P is a point of a region R. By Theorem 27, there exists in R a domain U containing P such that the complement of $\bar{U}$ is a connected set C. By Proposition 5, there exists a continuum T lying in U and separating P from a point of C. Let D_1 denote the complementary domain of T which contains P. Then D_1 is a subset of U and, consequently, of R. The boundary of D_1 is a subset of the continuum T which lies in $R - D_1$. Hence the space satisfies Axiom 5_1^*.

THEOREM 33. *Any locally divisible metric space S satisfying Moore's Axioms* 0–4 *in which Proposition* 5 *holds true is a subset of a plane or a sphere such that, if P is a point of a region R in S, there exists in R a connected domain D containing P whose boundary is either a simple closed curve or a simple open curve, and $S - \bar{D}$ is connected.*

Theorem 33 follows from the preceding theorem and Theorems 14, 33, 41 and 82 of the author's paper on topologically flat spaces.[24]

In order to settle numerous questions that arise in connection with the results of this section, the reader should note the following four facts concerning *metric spaces S which satisfy Moore's Axioms* 0–4. For each i $(1 \leqq i \leqq 6)$, let α_i denote an unbroken subsequence of α from Proposition 1 to Proposition i. For example, α_2 is Propositions 1, 2a, 2b, and 2c.

(1) In Examples 1–7, S is a subset of a plane, S is locally remotely connected, and either none of the propositions of α or only those of α_i $(1 \leqq i \leqq 5)$ are true in S, but Axiom $\bar{5}$ is false in S.

(2) An unpublished example of R. L. Moore's shows that there exists a space S in which Axiom $\bar{5}$ and all of the propositions of α hold true, but which is neither locally divisible nor a subset of a plane or a sphere.

(3) In Examples 8 and 9, S is a subset of a plane and Axiom $\bar{5}$ holds true in S but neither is S locally divisible nor do any of the propositions of α hold true in S.

(4) Remove from Example 4 the points of the X-axis whose abscissas are $\pm n^{-1}$ $(n = 1, 2, 3, \cdots)$ and all points whose ordinates are one or more. The subset remaining is a space S in the plane in which the propositions of α_5 hold true but S is not locally divisible and Axiom $\bar{5}$ is not true in S.

THE UNIVERSITY OF TEXAS.

[22] J. H. Roberts, *Concerning compact continua in certain spaces of R. L. Moore*, Bull. Amer. Math. Soc., vol. 39(1933), pp. 615–621.

[23] See the reference of footnote 7.

[24] Loc. cit.

THE MAPS DETERMINED BY THE PRINCIPAL CURVES ASSOCIATED WITH FIVE AND SIX POINTS IN THE PLANE

By C. Ronald Cassity

1. Introduction

Any set of n generally chosen points of the plane determines a set of principal curves (P-curves). This set is finite in number if $n < 9$. The P-curves associated with the points divide the projective plane into regions and thus form a *map* on the plane. Such a map will be described by telling what kinds of regions it contains, how many of each kind, and in what manner they fit together. Two maps will be considered as identical if and only if the description of one is also a description of the other.

The formation of a map by the division of the projective plane into twelve five-sided regions by the P-curves attached to four generally chosen points in the plane has been studied by Coble and Brahana [1].[1] The group associated with the figure is derived and the geometrical interpretation of its various types of elements discussed. The order of the group is 120. Each of its elements can be expressed algebraically, for they are either collineations or quadratic transformations. The map has the maximum possible amount of *regularity*, for every element of the group can be interpreted as a topological transformation of the map into itself.

The cases $n = 5$ and $n = 6$ lead, by mappings by the linear system of cubic curves on the base points, to the rational surface of order four in S_4 and to the general cubic surface in S_3, respectively. The P-curves associated with the points map into lines on the surfaces. Hence the maps determined by the P-curves in the plane are identical with those determined by the lines on the surfaces. It is the author's purpose to discuss these maps in detail. In particular, it will be shown that, whereas the map determined by four points has *complete* regularity, that determined by five points has only *partial* regularity and that determined by six points has *no* regularity. The total lack of regularity in the map on the cubic surface may be somewhat surprising to those familiar with the results obtained by Klein ([8], p. 570) and Zeuthen ([12], p. 8). Their descriptions of the map are correct, but are incomplete according to the definition given above.

§3.7 is devoted to the correction of an error made by Zeuthen ([12], p. 8).

2. The case $n = 5$

Five points of the plane are, for present purposes, "generally chosen" if no two are coincident and no three collinear. Associated with such a set of five

Received October 4, 1938.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

points are sixteen P-curves: (1) the conic, C, on the five points; (2) the ten lines ij $(i, j = 1, \cdots, 5; i \neq j)$; and (3) the directions at each of the five points P_i $(i = 1, \cdots, 5)$. In order to emphasize these latter as P-curves, each point will be represented by a small circle. One may think of punching a circular hole in the plane and then joining opposite points on the circle, thus introducing a figure like a Möbius band for each base point. The connectivity of surfaces as obtained from this idea agrees with that found by other means.

Under any element of a group, G_{1920}, of order $5! \times 16 = 1920$, the chosen set of base points passes into a *congruent* [5] set, and the P-curves of one set pass into the P-curves of the other set. The rational surface in S_4 is unaltered by passing from one set of base points to a congruent set. The elements of the group may be generated from the elements of the symmetric group, Σ_{120}, of degree five which permutes the base points in all possible ways and the Cremona transformations of orders two and three which have the given points as base points. A_{ijk} will be used to denote any quadratic transformation having i, j, k as self-corresponding fundamental points; A_{i^2jklm} to denote any cubic Cremona transformation having i as a double base point and j, k, l, m as self-corresponding simple base points. Either of these followed by a permutation will mean the Cremona transformation followed by the indicated permutation on the base points. In particular, $A_{ijk}(lm)$ represents the involutorial quadratic transformation which is determined by having i, j, k as self-corresponding base points and l, m as a pair of corresponding ordinary points.

Since any real conic can be projected into a circle [9], we shall (for convenience, not of necessity) assume the five base points to be chosen on a circle.

It is intuitively evident that *five points on a circle can give only one kind of a map*, i.e., the nature of the map is not changed by any motion of the points along the circle. For, as one point moves along the segment of the circle which connects the two adjacent points, the four lines from this point to the other four have intersections with the remaining six lines and directions at the five points whose relative positions with respect to other such intersections and directions are preserved. This map contains 16 pentagonal regions and 20 quadrangular regions. Four regions come together at each vertex. Adjacent to every side of any region is a region of different type. If the points be so numbered that they lie in the order 1, 2, 3, 4, 5 (henceforth called the natural order) on C, the map is that shown in Figure 1. It is a *semi-regular* map in the sense of Coxeter [6] who considers the converse problem of representing given groups by geometric figures. He considers his surfaces to be two-sided while our surface is one-sided, i.e., one of his k-sided regions can be applied to any k-sided region in k different ways while, on our surface, any k-sided region can be applied to any k-sided region in $2k$ ways, the additional factor of 2 being due to the possibility of "turning the polygon over" which can always be done by transporting it along a suitable path on a one-sided surface. Coxeter's semi-regular maps represent groups which, for Figure 1, would be of order 80. Actually the group is of order 160. It will be seen that *this is a subgroup of index* 12 *in the whole group*, G_{1920}.

Five elements in a cycle may have twelve different orders. It is easily seen that the order of the five elements on any one of the sixteen P-curves associated

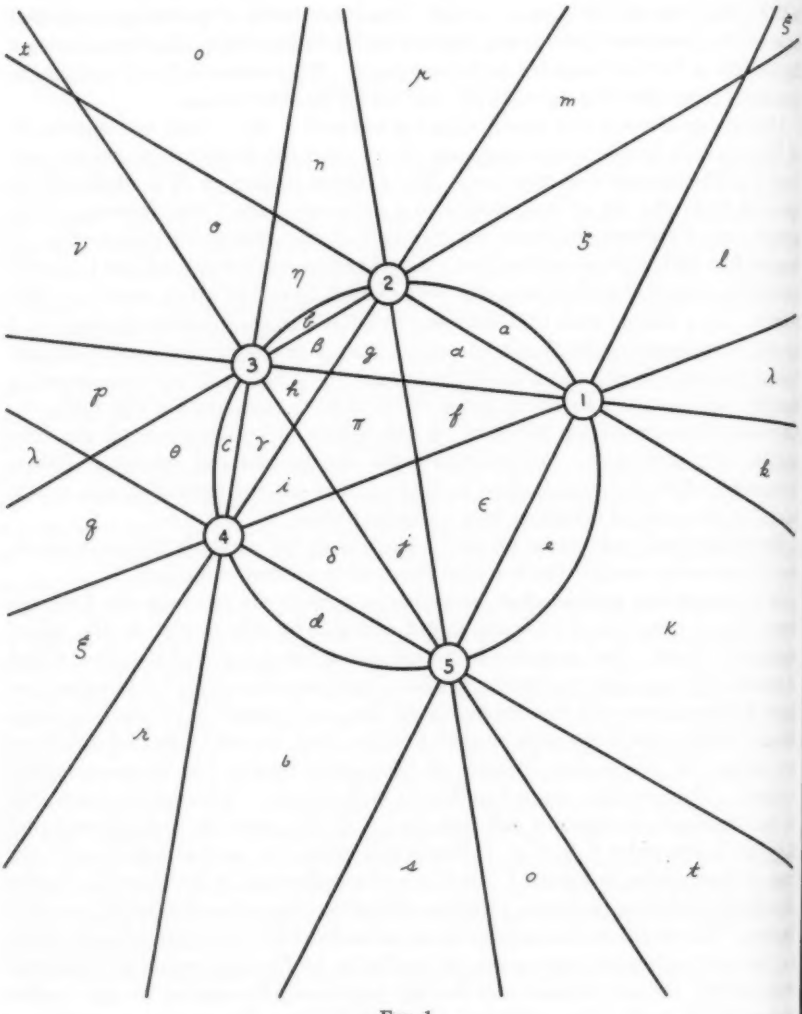

FIG. 1

with the case $n = 5$ determines the order of the five elements on each of the remaining P-curves. It is also evident that, under the permutations of Σ_{15}, the elements on any P-curve can be made to lie in any of the twelve possible

orders. For the given set of five points this amounts merely to a renaming. Therefore, the subgroup of G_{1920} which leaves any prescribed order invariant will be of index 12 in G_{1920}. That *this group, G_{160}, coincides with the previously-mentioned group of order* 160 is proved by observing that, under the choice of the natural order for the points on C, both groups may be generated by the elements (12345), (12)(35), $A_{123}(45)$. The elements (12345) and $A_{123}(45)$ generate a group of order 80 which is the group of the map in the sense of Coxeter. The element (12)(35) "turns over" the regions marked ξ, i, π, α, a, ζ in Figure 1.

The elements of G_{1920} all leave the rational quartic surface in S_4 invariant. They really represent the sixteen possible ways of selecting a line on the surface and the 120 possible ways of naming the five lines which are met by this line. In the plane representation the isolated line may be taken as C. The five lines which meet it are then represented by $P_1, \cdots, P_5$ and the 120 ways of naming these are the elements of Σ_{120}.

If we start with any definite naming of the lines, say that indicated in Figure 1, some of the other namings can be considered as obtainable from the first one by topological transformations of the surface into itself. These, naturally, form a subgroup of G_{1920}. As previously indicated, this subgroup is of order 160 and is one of a conjugate set of twelve such subgroups, each being characterized geometrically as leaving invariant the cyclic order of the intersections of every line with the five others which it meets. Since only part of the elements of G_{1920} can be considered as topological transformations of the surface into itself, the map on the surface is said to have only *partial* regularity.

The cross-cut of the twelve subgroups mentioned above is an invariant subgroup, G_{16}, whose elements are all of period two. They are the ten elements of the type $A_{ijk}(lm)$, five of the type A_{i^2jklm}, and the identity. This group is regular on the sixteen lines.

The G_{1920} is the Galois group of the algebraic problem involved in the determination of the 16 lines of an F^4 in S_4, the complete intersection of two quadrics. *The G_{16} contains the only elements of G_{1920} which, in the plane or in S_4, are algebraic.*

3. The case $n = 6$

3.1. **The nature of the map.** It is well known that the cubic curves on six generally chosen points map the plane into a general cubic surface and, conversely, that the general cubic surface can be mapped on the plane, the twenty-seven lines of the surface mapping into the twenty-seven P-curves associated with six points of the plane. The P-curves are: (1) the directions at each of the points P_i $(i = 1, \cdots, 6)$, (2) the conics C_i $(i = 1, \cdots, 6)$, where the notation is such that C_i is on all the points except P_i, and (3) the fifteen lines ij $(i, j = 1, \cdots, 6;\ i \neq j)$. The group of the surface is of order 51840. It may be generated from the 720 elements of the symmetric group of degree six and the 71 Cremona transformations of orders 2, 3, 4, and 5 determined by the six base points.

The results obtained by Klein [8] and Zeuthen [12] will be assumed. They

showed that the involutions set up on the lines of the surface by the residual conics cut out by the pencils of planes through the lines are such that 15 have real double points and 12 have imaginary. The number and types of regions were shown to be as follows: (A) ten triangles whose thirty bounding line-segments are the thirty segments containing the double points of the involutions of the 15 lines on which the involutions have real double points; (B) thirty pentagons each having one side in common with one of the ten triangles; (C) thirty quadrangular regions each having two adjacent sides in common with two of the pentagons which are adjacent to a triangle; (D) sixty other quadrangular regions each having two adjacent sides in common with regions of type (B) or (C) and the other two sides in common with other regions of type (D).

It is possible to construct several different maps answering all of the description above. I shall show in §§3.2-3.5 that *only one of these can exist as the map on a cubic surface.*

This paper deals only with the most general real cubic surface having 27 real lines.

In order to make certain that the surface under consideration is general in the sense that it have no double points, it is necessary and sufficient that, in the plane representation, no two points be coincident, no three collinear, and the six not on a conic. These are equivalent conditions under the group of order 51840, for a set of points satisfying one of the conditions can be transformed into a congruent set satisfying any one of the other conditions.

In order that none of the ten triangles of the general map may shrink to a point, it is necessary and sufficient that, in the plane representation, no set of three lines ij, kl, mn be concurrent and that no line ij be tangent to the conic C_i. These also are equivalent conditions under the group.

3.2. **A canonical position of the base points.** Both Klein and Zeuthen noticed that the twelve lines on which the involutions have imaginary double points form a double six. When this double six is taken as the fundamental one for the plane representation of the surface, the ten triangles of the map are bounded entirely by the fifteen lines denoted by ij. I shall give conditions on the base points which cause the triangles to be bounded entirely by the lines of the plane representation.

Two auxiliary theorems will be needed. Since these are so simple as to be almost obvious, they will be given without proofs.

THEOREM I. *If the five base points on C_n lie in the order i, j, k, l, m, then the orders of the lines in the pencils on the base points are ij, ik, il, im; ji, jk, jl, jm; ki, kj, kl, km; li, lj, lk, lm; mi, mj, mk, ml; and conversely.*

THEOREM II. *In the five-point map there is one and just one such pentagonal region as is marked π in Figure* 1. *If the base points lie in the order i, j, k, l, m on C, this region is bounded by the lines ik, jl, km, li, mj.*

Schur [9] remarked that a theorem like Theorem I might be given, but I have been unable to find any such theorem in the literature.

THEOREM III. *A necessary and sufficient condition that the six points* P_i $(i = 1, \cdots, 6)$ *be the base points of a plane representation corresponding to the fundamental double six on which the involutions have imaginary double points is that each point* P_i *lie within the region* π_i *formed for the conic* C_i .

Wieleitner [11] says, without proof, that if P_i lies within π_i for one value of i, then the same is true for all values of i. This may be proved by using Theorem I.

The proof of Theorem III, then, is reduced to the problem of showing that the six points will have the property required if and only if one point, say P_6, lies within the proper pentagon (π_6). If Figure 1 and Theorem I are used, it will be seen that the type of plane configuration obtained depends on the region of Figure 1 in which P_6 is placed. If, e.g., P_6 be placed in any of the regions a, b, c, d, e, the representations obtained will be similar, and it is only necessary to consider one region, say a. If P_6 be placed in a, the conic C_1 will necessarily run through the regions a, η, c, ι, s (or e), ϵ, f, α, a. Thus, in either s or e, there will be a triangle bounded by P_5 , C_1 , 15, and the triangles will not be bounded entirely by lines of the type ij . In a similar way it may be seen that no region except π will satisfy the requirement. That the condition is satisfied by placing P_6 in π_6 is shown by Figure 2.

3.3. **A topological representation of the map.** Figure 2 is a plane representation of the map. It is similar to the incomplete figure drawn by Wieleitner [11] at the end of his book. Figure 3 is a topological representation of the map in the finite portion of the plane. It shows up clearly the different types of regions found by Zeuthen ([12], p. 8). In this figure, the outside boundaries must be considered as joined together in the manner implied by their names as is usual in such topological figures.

3.4. **The uniqueness of the regions.** The apparent symmetry (regularity) in Figure 3 is a bit misleading. *There is not a single one of the* 130 *regions which cannot be distinguished from all the others.* We shall begin with the triangles.

Every triangle is bounded by three pentagons. Every pentagon is bounded by a triangle and four quadrangular regions. Some pentagons, however, are "vertical to" two other pentagons while others are vertical to three. If a pentagon is vertical to only two other pentagons (the two adjacent to the same triangle), then the vertex opposite the triangle is vertical to a quadrangle, and the opposite vertex of the quadrangle is vertical to another triangle. Any triangle may be distinguished by the number of times that one may "go out" from it to another triangle via two vertical pentagons and the number of times that one may "get in" to it from another triangle via a vertical pentagon and quadrangle. The two numbers in each triangle of Figure 3 are these two numbers, respectively.

It will be seen that there are 20 possible arrangements of the character mentioned: one each of the types 33, 32, 31, 30, 23, 20, 13, 10, 03, 02, 01, 00 and two each (one symmetric and one non-symmetric) of the types 22, 21, 12, 11.

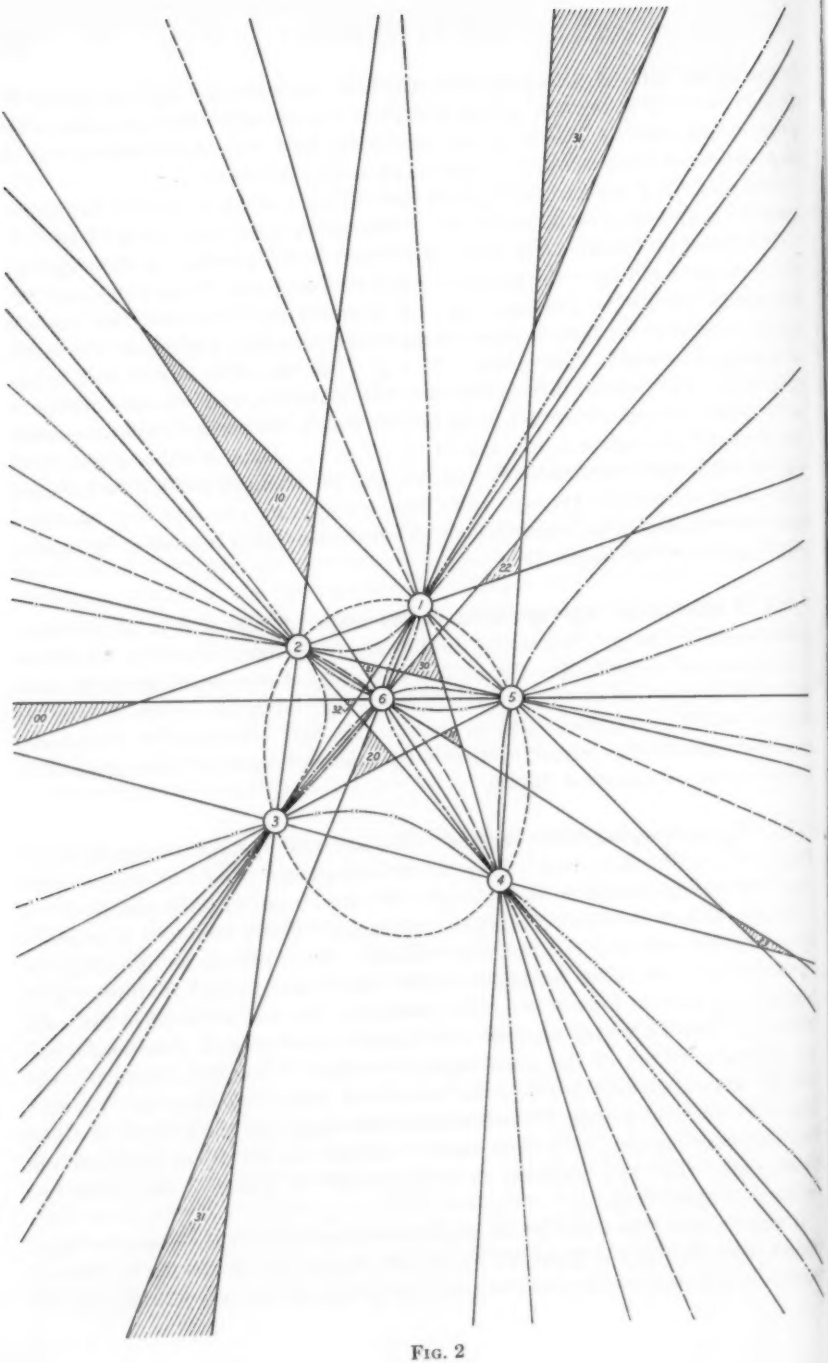

FIG. 2

Of these possible types not more than ten can occur on any particular surface since the map on the surface has only ten triangles. In Figure 3 these are the types 33, 32, 31, 30, 22 (symmetric), 21 (non-symmetric), 20, 11 (symmetric), 10, and 00. It will be proved in §3.5 that every general cubic surface with 27 real lines is of this type.

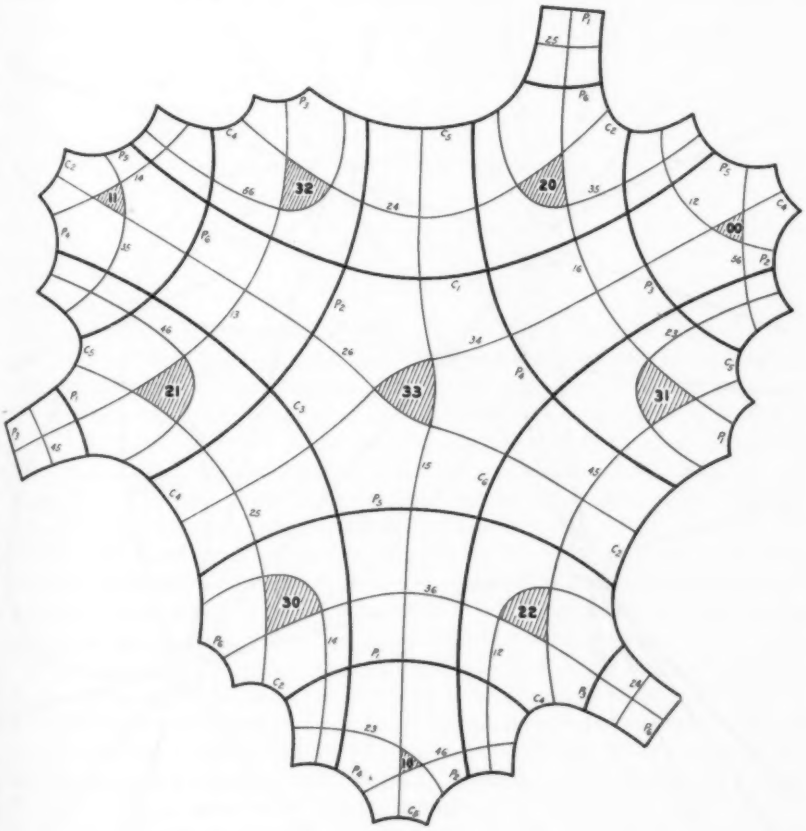

FIG. 3

Since the triangles are all different, each one of the fifteen lines ij may be distinguished by telling which two triangles it bounds. The names of the remaining twelve lines are then unique except for the possibility of interchanging P_i with C_i, i.e., applying the quintic Cremona transformation which interchanges the two sixers of the fundamental double six. Which of these is to be used may be determined by giving the boundaries of any pentagonal region. Then every region of the map has a unique set of bounding lines and is, therefore, distinguished from all the others.

A connection between the naming of the lines and the well known functions of Joubert will be discussed in §3.6.

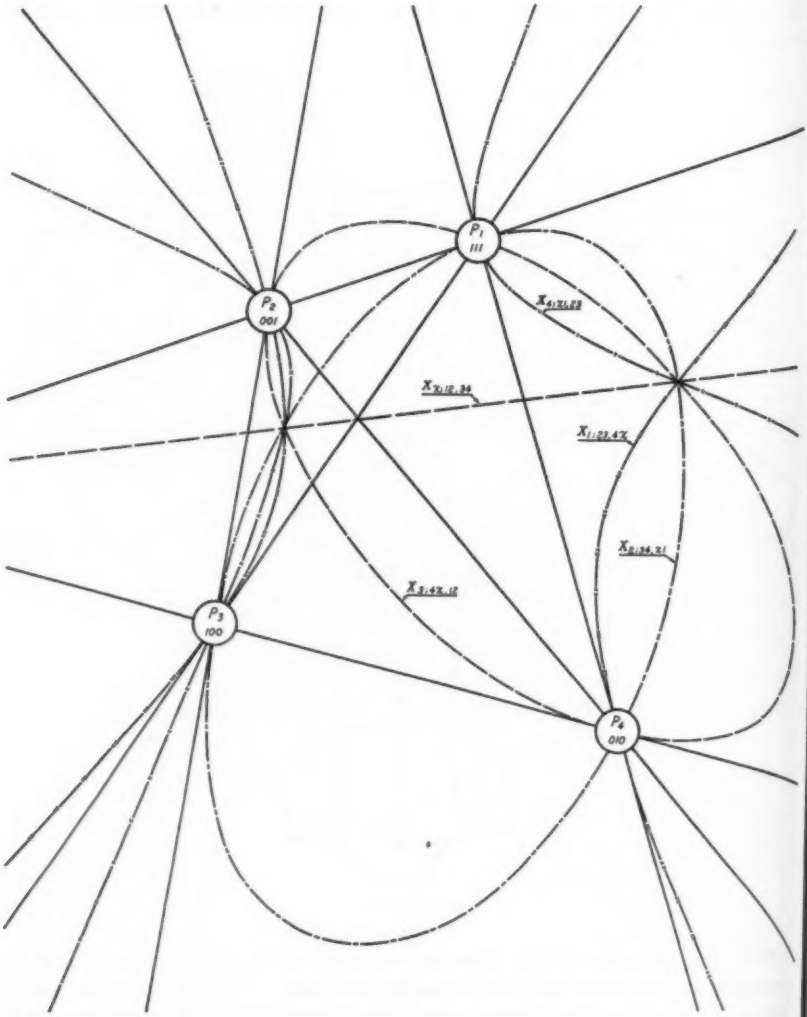

FIG. 4

3.5. **The maps are all identical.** Let any one conic, say C_6, be chosen. If it is not an ellipse, let it be projected into one. Of course, the whole plane configuration must be projected along with C_6. On C_6 are the five points $P_1, \cdots,$

P_5. Let them be so numbered that they lie in the natural order on C_6. Let the coördinate system in the plane be so selected that P_1, P_2, P_3, P_4 have the coördinates (1, 1, 1), (0, 0, 1), (1, 0, 0), and (0, 1, 0), respectively. Then P_5 must lie in the region bounded by P_1, 14, P_4, 43, 21, for otherwise the order on C_6 could not be the natural order.

Let $X_{i;jk,lm} = 0$ be the condition that the point i be collinear with the diagonal points jk, lm and jl, km of the complete quadrangle on the points j, k, l, m. Then the region in which P_5 must lie may be divided into ten subregions, Figure 4, by the following five loci:

$$X_{x;12,34} \equiv x_0 - x_1 + x_2 = 0,$$

$$X_{1;23,4x} \equiv x_0x_1 - x_0x_2 - x_1x_2 = 0,$$

$$X_{2;34,x1} \equiv x_2^2 + x_0x_1 - 2x_1x_2 = 0,$$

$$X_{3;4x,12} \equiv x_0^2 - x_1x_2 = 0,$$

$$X_{4;x1,23} \equiv x_1^2 + x_0x_2 - 2x_0x_1 = 0.$$

These are conjugate loci under the group of order 120 on the four points $P_1, \cdots, P_4$ [4]. The complete set of conjugates consists of the three sides of the diagonal triangle and eight other conics, but these are not necessary for present purposes.

These five loci are concurrent at the two points $(1 + \sqrt{5}, 3 + \sqrt{5}, 2)$ and $(1 - \sqrt{5}, 3 - \sqrt{5}, 2)$ of which only the latter lies in the region under consideration. They are two of the twelve points discussed by Clebsch [3] and Coble [4]. The twelve points lie by fours on the three sides of the diagonal triangle of the complete quadrangle on the points $P_1, \cdots, P_4$. They may be obtained by pairs by considerations exactly similar to those above only using different orderings for the points in the quadrangles. The coördinates are the possible permutations of those given above. One point lies in each of the twelve pentagonal regions of the map determined by the four points $P_1, \cdots, P_4$.

If P_5 be placed on any $X_{i;jk,lm}$, then the line $i6$ may not be made to cross one of the intersections jk, lm or jl, km without crossing the other. By moving P_5 to one side or the other of $X_{i;jk,lm}$, the line $i6$ may be moved across either of the intersections without crossing the other. At any point within any of the ten subregions these relations are all the same. Nothing is to be gained by placing P_5 on any $X_{i;jk,lm}$ since we are studying maps which have ten triangles none of which degenerates to a point. Hence there are ten possible types of figures under the present classification. We show next that these types are not different, but are, for present purposes, all alike; cf. Figure 5.

Ten figures having been drawn, one having P_5 in each of the ten subregions of Figure 4, and then the figures completed according to the method to follow, it was found that no one of the ten could be distinguished from any other.

The points $P_1, \cdots, P_5$ having been chosen (say on an ellipse) and in their natural order, the point P_6 must be placed within the pentagon π_6. This pentagon may be subdivided into at most forty-one subregions by the five pairs

of lines which go through π_6 and join each of the points $P_1, \cdots, P_5$ to two diagonal points of the complete quadrangle on the other four points. These ten lines form a Desargues Configuration (Figure 5) and divide π_6 into a set of subregions which are such that, as P_6 is moved across one of the lines from P_i, one of

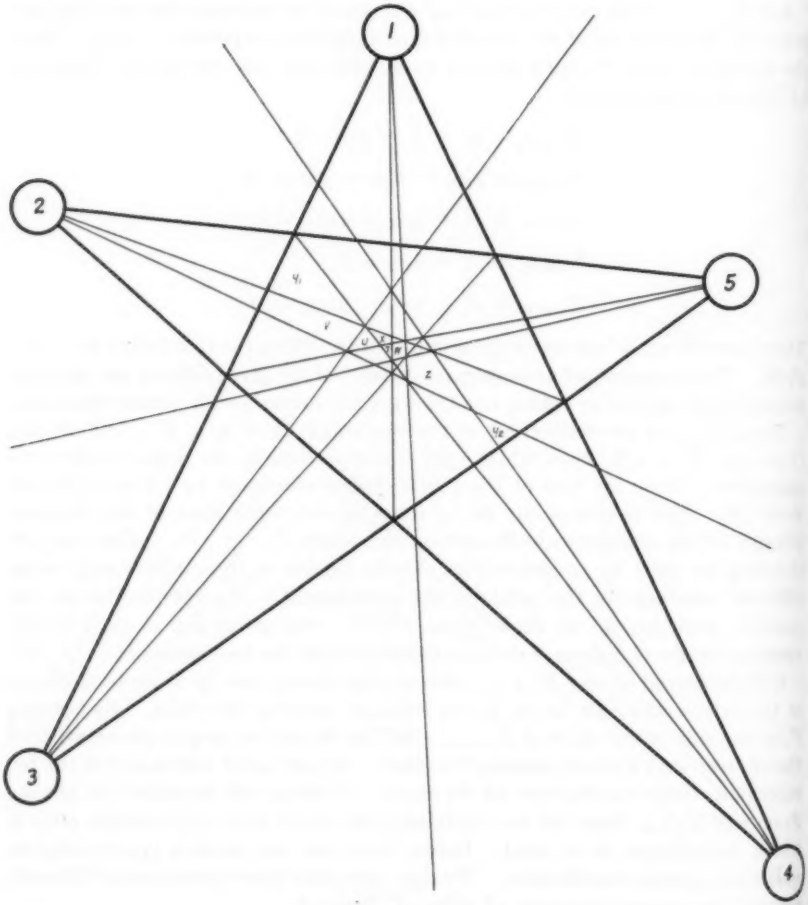

FIG. 5

the triangles which is bounded by $i6$ shrinks to a point and reappears again on the opposite side of the line $i6$.

It is fairly easy to convince oneself that the configuration divides the projective plane into 36 regions. Five of these which lie in the infinite portion of the metric plane may have two distinct portions of each within π_6, thus making 41 possible

subregions in π_6. The positions of $P_1, \cdots, P_5$ may be such that some of these possible subregions lie partly or entirely outside π_6. This will not affect our argument since we are only interested in showing that such cases as are possible are all of the same kind.

In trying to find some unique method of naming the subregions of π_6, it is natural to begin with that one subregion which lies between all the pairs of lines drawn from the points $P_1, \cdots, P_5$. This region is a quadrangle having three points of the configuration as its vertices. Proceeding across the fourth vertex, one arrives at the subregions marked w, x, y_1 in Figure 5. Coming in from the other side are the subregions y_2, z, w. It may be observed that the subregions have complete symmetry with respect to this "axis of symmetry". The ten figures drawn for the ten subregions of Figure 4 and completed by the Desargues Configurations all have exactly this description, i.e., they are, for our purpose, identical.

There remains only to examine the possibility of obtaining different maps as P_6 is made to lie in the 41 subregions of π_6.

As a matter of convenience, some terminology will be introduced first. The words to be used are suggested by a study of Figure 3.

A triangle with its three adjacent pentagons and the three quadrangles each having a vertex in common with the triangle may be thought of as combined into a single region and called a *fundamental hexagon*. The four quadrangles of type (D) having a vertex in common may be combined into a single region and called a *fundamental quadrangle*. Then two fundamental hexagons having a vertex in common will be called *vertical* and the same term applied to the corresponding triangles. Two fundamental hexagons which are not vertical are necessarily adjacent to opposite sides of a fundamental quadrangle. Such fundamental hexagons and their corresponding triangles are *separated*.

If any cubic surface contains triangles of the types 33 and 00, these triangles must obviously be separated. Every other triangle is vertical to one or the other of two separated triangles. Hence any map containing triangles of the types 33 and 00 is completely determined and must be exactly that shown in Figure 3. Hence we proceed to show that every map contains these two types. This will be done by tabulating the numbers pq characterizing the triangular regions. This tabulation is started in Table II. It may be completed quickly by using Table I which is easily obtained from Figure 3. Table I is a schematic arrangement representing four of the triangular regions as they appear in Figure 3. The six numbers surrounding another number are the characteristic numbers of the six triangles which are vertical to the triangle whose characteristic numbers appear in the center. The underlined numbers represent the corresponding topological characteristics of the triangles after the shift of one line across the central triangle. It must be remarked that the clockwise or counterclockwise order of the six others around any one is quite immaterial since the surface is one-sided and, viewed from one side, the order is the reverse of that observed from the other side.

TABLE I

21	30		20	11		00	22		21	10	
31	20		10	21		10	21		20	11	
32	33	22	32	22	33	32	11	31	30	00	31
22	33	32	33	22	32	31	11	32	31	00	30
30	21		11	20		22	00		10	21	
20	31		21	10		21	10		11	20	

TABLE II

Region for P_6 / Triangle	w	t	x	u	v	y_1	⋯
12, 34, 56	11	10	10	00	00	10	⋯
23, 45, 16	22	22	32	31	31	32	⋯
34, 51, 26	20	21	31	32	33	33	⋯
45, 12, 36	30	31	21	22↔22		21	⋯
51, 23, 46	31	30	20	20	10	00	⋯
13, 24, 56	33	33↔33		33	32	31	⋯
24, 35, 16	00↔00		00	10	20	20	⋯
35, 41, 26	32	32	22	21	11↔11		⋯
41, 52, 36	21	20	30	30	30	30	⋯
52, 13, 46	10	11	11↔11		21	22	⋯

The above table is to be read as follows: As the point P_6 is moved across the line 1; 24, 35 from the region w to the region t of Figure 5, the triangle which has the bounding lines 24, 35, 16 and the characteristic numbers 00 shrinks to a point and reappears again on the opposite side of each of its bounding lines. No shift is made across any other triangle. The triangles originally having the characteristic numbers given in column one will, after the shift, have those given in column two; etc. to the other columns.

The only arrangements not altered by a shift are those around the four triangles having the symmetric numbers 33, 22, 11, 00. These are the ones given in Table I. In completing Table II, the only triangles across which lines are moved are those same four. Thus the triangles always have the numbers 33, 32, 31, 30, 22, 21, 20, 11, 10, 00. We conclude, then, that *the map on every general cubic surface with* 27 *real lines is that shown in Figure* 3.

3.6. **The group of the map.** In fitting a map topologically upon itself, regions of a given number of sides must fit upon regions having that same number of sides. The adjacent regions must fit upon the adjacent regions, etc. In particular, every triangle must fit upon a triangle. The adjacent pentagons and quadrangles must fit upon the adjacent pentagons and quadrangles, etc. Pro-

ceeding outward in this manner from any triangle of the map, we see that every triangle must be fitted upon a triangle having the same characteristic numbers, i.e., upon itself. Thus the whole map may be fitted upon itself in only the identical way. Hence *the group of the map is the identity, and the map has no regularity whatever.*

If we look at this fact from a different viewpoint, it is clear that it must be possible to start with the whole group, G_{51840}, of the surface and reduce it, by successive steps, to the group of the map.

This reduction may be carried out by first selecting the unique double six which has no line bounding any of the ten triangles of the map. Since there are thirty-six double sixes, the group will be reduced to one of index 36, i.e., to a G_{1440}. The elements of this group are the 720 elements of the symmetric group of degree six and their products by $A_{i^2j^2k^2l^2m^2n^2}$, the quintic Cremona transformation having the six base points of the plane representation as double principal points. The quintic transformation interchanges the two halves of the fundamental double six, but leaves the other lines invariant. The permutations of Σ_{720} permute the six lines of either sixer in all possible ways and those of the other sixer in the same way. They also permute the remaining 15 lines, ij, but in a less obvious way.

The 15 lines ij can be grouped into 15 sets of three: ij, kl, mn. Each of these sets represents one of the 45 tritangent planes of the surface. Five of these 15 are different from the remaining ten since only ten of the sets can form boundaries of the ten triangles of the map. Any two sets have only one line in common. The six functions of Joubert [7] make exactly this distinction between the lines and the sets of three lines. I shall use these functions in a form similar to that in which they were written by Coble ([5], I, p. 167):

$$A = 25, 13, 46 + \underline{51, 42, 36} + 14, 35, 26 + 43, 21, 56 + 32, 54, 16,$$

$$B = \underline{53, 12, 46} + \underline{14, 23, 56} + \underline{25, 34, 16} + \underline{31, 45, 26} + \underline{42, 51, 36},$$

$$C = 53, 41, 26 + \underline{34, 25, 16} + 42, 13, 56 + 21, 54, 36 + 15, 32, 46,$$

$$D = \underline{45, 31, 26} + 53, 24, 16 + 41, 25, 36 + 32, 15, 46 + 21, 43, 56,$$

$$E = 31, 24, 56 + \underline{12, 53, 46} + 25, 41, 36 + 54, 32, 16 + 43, 15, 26,$$

$$F = 42, 35, 16 + \underline{23, 14, 56} + 31, 52, 46 + 15, 43, 26 + 54, 21, 36.$$

Any one of these functions (B will be used in order to agree with Table II) isolates five sets of three lines. These may be taken as the five tritangent planes which do not cut out triangles of the map. The remaining ten sets, then, give the boundaries of the triangles of the map. Hence, each of the six functions determines a naming for the triangles and, conversely, any naming of the triangles determines one of the functions. Therefore, within the G_{1440}, there are six ways of naming the triangles. One of these ways having been chosen, the group is reduced to a subgroup of index six, i.e., to a G_{240}. The elements of G_{240} are the 120 elements [2, 10] of one of the six subgroups of Σ_{720} which are of

degree six and their products with $A_{i^2j^2k^2l^2m^2n^2}$. Thus it might be said that the elements of Σ_{720} permute the ten sets of lines ij, kl, mn through six conjugate sets which correspond to the six functions of Joubert.

The five different terms underlined above are the five tritangent planes which do not cut out triangles of the map as they are written in Table II. The ten triangles might equally well be named by the pairs of functions in which they lie. Thus, having isolated B, the triangles have two equivalent sets of names:

12, 34, 56 = AD,	13, 24, 56 = CE,
23, 45, 16 = AE,	24, 35, 16 = DF,
34, 51, 26 = EF,	35, 41, 26 = AC,
45, 12, 36 = CF,	41, 52, 36 = DE,
51, 23, 46 = CD,	52, 13, 46 = AF.

From arguments similar to those used in §3.2, it may be seen that the above selection of the sets of lines bounding triangles requires that the points lie in one of two possible orders on the conics which they determine. For example, the above selection requires that the order on C_6 be either 1, 2, 3, 4, 5 or (its square) 1, 3, 5, 2, 4, and the corresponding orders on the other C_i.

In order to reduce the group to still lower order, it is necessary to use the properties derived in §§3.4 and 3.5. Since every triangle must fit upon itself, the line which joins two separated triangles must fit upon itself. Every one of the 15 lines ij joins two separated triangles and must, therefore, fit upon itself. The intersection of the fixed lines ij and ik must, then, be the fixed point P_i. This reduces the group to order 2: $I, A_{i^2j^2k^2l^2m^2n^2}$. But the quintic transformation destroys the identity of the regions of types (B), (C) and (D). Hence, we conclude, as before, that *the group of the map is the identity.*

3.7. **The shifts of Zeuthen.** Consider the two line-segments which form the common boundary of an adjacent fundamental hexagon and fundamental quadrangle. One of the segments bounds a region of type (B) and a region of type (D). The other segment bounds regions of types (C) and (D). Let the first of these be called a segment d and the second a segment e. Zeuthen ([12], p. 8) says that "the order of these two segments—or that of the regions of types (B) and (C) which they bound—can vary independently of that of the other segments d and e of the same line, as may be seen by varying the size of the ovals of the parabolic curve. At the moment when a change is made in the order, one of the triangles shrinks to a point."

The idea of changing the order of the regions of types (B) and (C) of a fundamental hexagon by "shifting" a line across the central triangle appeared in §3.5. The shifts, however, appeared as incidental consequences of moving the positions of the base points of the plane representation, and no such statement as the above was either made or implied.

The fundamental hexagons 00, 11, 22, 33, 30 lie along a line (P_5 in Figure 3). The order of the segments d and e along this line is $e, d; e, d; e, d; e, d; d, e$. Zeuthen's statement requires that the segments around 30 could be interchanged leaving the order $e, d; e, d; e, d; e, d; e, d$. But it has been proved in §3.5 that the map on every general cubic surface is that shown in Figure 3. This map contains no line on which the order of the segments d and e is $e, d; e, d; e, d; e, d; e, d$. Hence *Zeuthen's statement is incorrect.* It may be replaced by Theorems IV and V which are the answers to the following two questions:

(1) Across which triangles, as given by their characteristic numbers, may lines be shifted without first making other shifts at other portions of the surface?

(2) If a line may not be shifted across a given triangle without first making other shifts, how many other shifts must be made first and in how many ways may they be selected?

It appeared in §3.5 that the map on every general cubic surface has ten triangles whose characteristic numbers are 33, 32, 31, 30, 22, 21, 20, 11, 10, 00, i.e., pq ($p, q = 3, 2, 1, 0; p \geqq q$). A little reflection will show that a shift of any one of its three bounding lines across the triangle pq changes that triangle's numbers to qp. Since $p \geqq q$ for every triangle of the map, it is not possible to shift a line across any triangle except those having the symmetric numbers given in Table I. In completing Table II, shifts are made across all four of these triangles. This leads to the question of whether $p = q$ is also a sufficient condition for the shifting of a line across the triangle without first making other shifts.

The subregions of Figure 5 are regions of two, three or four sides (counting only the lines of the Desargues Configuration as sides). If P_6 be placed in a subregion of four sides, say in w or x, then a line can be shifted across all four of the symmetric triangles by moving P_6 only. If P_6 be placed in a two-sided or three-sided region, then only two or three of the shifts may be made by moving only P_6. There remains the possibility of moving both P_5 and P_6.

In order to check the remainder of this proof, one will find it practically a necessity to have all ten figures like Figure 5. I have checked the statements to follow for all 41 subregions. Examples will be given to show the typical line of argument, for a complete enumeration would be long and probably useless.

THEOREM IV. *A necessary and sufficient condition that a line can be shifted across a triangle without making any other shifts is that the triangle have one of the symmetric characteristic numbers.*

The necessity has already been shown. As an example of the demonstration of the sufficiency, suppose P_6 to be placed in the subregion t of Figure 5. This is a three-sided region. By moving P_6 across the lines 1;24,35, 4;13,25 and 5;13,24, a line may be shifted across the triangles 00, 11 and 33, respectively. In order to shift a line across the triangle 22, it would be necessary to move P_6 across 1;23,45. This may be done by moving both P_5 and P_6. First move P_5 across $X_{1;23,4z}$ (my Figure 5 is drawn with P_5 in the region bounded by P_4, 14, $X_{z;12,34}$ and $X_{1;23,4z}$ of Figure 4), moving P_6 at the same time, if necessary, so that

no line of the Desargues Configuration crosses P_6. This has the effect of replacing 1;24,35 by 1;23,45 as one of the boundaries of the subregion in which P_6 lies. P_6 can now be moved across 1;23,45, and a shift has been made over the desired triangle, 22. One must take care not to move P_5 in such a way that the subregion in which P_6 lies is destroyed. This also has been checked and found to be always possible.

Suppose it is desired to shift a line across a given triangle. If that triangle does not have a symmetric characteristic number, how many other shifts must be made before the desired one may be accomplished? Only shifts over triangles vertical to the given one will be counted, for the remaining three are separated from it and a shift across one of them does not alter the characteristic numbers of the one under consideration. A table giving the relations among the triangles follows. An x means that the triangles are separated. A 0 means that they are the same triangle. Those not marked are vertical. As a matter of possible interest, attention is drawn to the fact that the *separated* marks of this table are the *incidence* marks of the points and lines of a Desargues Configuration.

TABLE III

	00	10	11	20	21	22	30	31	32	33
33	x	x	x							0
32	x			x	x				0	
31		x		x		x		0		
30			x		x	x	0			
22	x					0	x	x		
21		x			0		x		x	
20			x	0				x	x	
11			0	x			x			x
10		0			x			x		x
00	0					x			x	x

Before a line can be shifted across any triangle, enough other shifts must be made to change that triangle's number into a symmetric number. Table I shows that a single shift changes at most one of the two numbers characterizing each triangle and that number only by 1. Hence, the triangle pq can not be changed into a symmetric one in less than $p - q$ shifts. The following is an example of the method which may be used to show that $p - q$ shifts are sufficient and that these may be chosen in $p - q + 1$ different ways. Consider the triangle 30. It is vertical to both 33 and 00. Let a shift be made across one of these, say 00. The new numbers for 30 are 31. 31 is vertical to 00, 11 and 33. Let a shift be made across one of these, say 11. (Do not use 00, for that changes 31 back to 30.) This changes 31 into 32. A further shift across 22 changes 32 into 33. This is the series of steps indicated in (1) below. The lines may now be shifted across 33, i.e., across the triangle whose original numbers were 30. Three other shifts have been made first. The four possible ways of choosing

the three triangles across which shifts must be made first are given in the following scheme:

$$(1)\quad 30 \xrightarrow[00]{} 31 \xrightarrow[11]{} 32 \xrightarrow[22]{} 33$$

$$(2)\quad \begin{cases} 30 \xrightarrow[00]{} 31 \xrightarrow[11]{} 32 \xrightarrow[33]{} 22 \\ 30 \xrightarrow[00]{} 31 \xrightarrow[33]{} 21 \xrightarrow[11]{} 22 \\ 30 \xrightarrow[33]{} 20 \xrightarrow[00]{} 21 \xrightarrow[11]{} 22 \end{cases}$$

$$(3)\quad \begin{cases} 30 \xrightarrow[00]{} 31 \xrightarrow[33]{} 21 \xrightarrow[22]{} 11 \\ 30 \xrightarrow[33]{} 20 \xrightarrow[22]{} 10 \xrightarrow[00]{} 11 \\ 30 \xrightarrow[33]{} 20 \xrightarrow[00]{} 21 \xrightarrow[22]{} 11 \end{cases}$$

$$(4)\quad 30 \xrightarrow[33]{} 20 \xrightarrow[22]{} 10 \xrightarrow[11]{} 00$$

The three ways given under (2) and (3) look somewhat different, but they represent shifts across the same three original triangles. They are the three originally numbered 00, 10 and 33 in (2) and those originally numbered 00, 32 and 33 in (3). Schemes like this one, though shorter, may be made for each of the other five non-symmetric characteristics. This leads to

THEOREM V. *In order that a line may be shifted across the triangle having the characteristic numbers* pq, $p - q$ *other shifts must be made first, and these may be chosen in* $p - q + 1$ *different ways.*

BIBLIOGRAPHY

1. H. R. BRAHANA and A. B. COBLE, *Maps of twelve countries with five sides with a group of order 120 containing an ikosahedral subgroup*, Amer. Jour. of Math., vol. 48(1926), pp. 1–20.
2. A. CAYLEY, *On the substitution groups for two, three, four, five, six, seven and eight letters*, Quarterly Jour. of Math., vol. 25(1891), pp. 71–88.
3. A. CLEBSCH, *Ueber das ebene Fünfeck*, Math. Ann., vol. 4(1871), pp. 476–489.
4. A. B. COBLE, *An application of the form-problems associated with certain Cremona groups to the solution of equations of higher degree*, Trans. Amer. Math. Soc., vol. 9(1908), pp. 396–424.
5. A. B. COBLE, *Point sets and allied Cremona groups*, I, Trans. Amer. Math. Soc., vol. 16(1915), pp. 155–198; II, Trans. Amer. Math. Soc., vol. 17(1916), pp. 345–385; III, Trans. Amer. Math. Soc., vol. 18(1917), pp. 331–372.
6. H. S. M. COXETER, *The groups* $G^{m,n,p}$. Will soon appear in Trans. Amer. Math. Soc.
7. P. JOUBERT, *Sur l'équation du sixième degré*, Comptes Rendus, vol. 64(1867), pp. 1025–1029 and 1081–1085.
8. F. KLEIN, *Ueber Flächen dritter Ordnung*, Math. Ann., vol. 6(1873), pp. 551–581.
9. F. SCHUR, *Ueber die Projection von fünf Punkten einer Ebene in fünf Punkte eines Kreises*, Zeit. für Math. und Phys., vol. 39(1894), pp. 247–249.
10. J. VERONESE, *Interprétations géométriques de la théorie des substitutions de n lettres, particulièrement pour* $n = 3, 4, 5, 6$, *en relation avec les groupes de l'hexagramme mystique*, Ann. di Mat., vol. 11(1883), pp. 93–236.
11. H. WIELEITNER, *Über die Flächen dritter Ordnung mit Ovalpunkten*, Speyer, 1901.
12. H. G. ZEUTHEN, *Études des propriétés de situation des surfaces cubiques*, Math. Ann., vol. 8(1875), pp. 1–30.

UNIVERSITY OF ILLINOIS.

FUNCTIONS WHICH ASSUME RATIONAL VALUES AT RATIONAL POINTS

By J. W. Green

1. **Introduction**. Of the continuous functions which assume rational values for rational values of the argument, the familiar examples either are extremely regular, as the rational or piecewise rational functions, or else exhibit some extremely irregular properties. For example, the functions $x(t)$, $y(t)$ defining Peano's area filling curve assume rational values for rational t and are nowhere derivable. Likewise the familiar function defined with respect to Cantor's ternary set as $\frac{1}{2}$ on the middle extracted one-third, $\frac{1}{4}$ and $\frac{3}{4}$ on the extracted middle one-thirds of the left- and right-hand remaining intervals, respectively, etc. can be shown to assume rational values for rational x. This function possesses a derivative at no point except points where it is piecewise rational.

It is desired, then, to investigate what kinds of functions may possess the property of assuming rational values at rational points, which property we shall denote by (A). For example, it might be asked whether the only analytic functions with property (A) are the rational functions, or it might be asked if it is possible that a function exist with property (A), the function being analytic in no interval and yet, say, having a continuous derivative.

The author has been informed that something of this nature was discussed in a conversation between Weierstrass and Hilbert and that Hilbert exhibited an example; however, no record of the conversation seems to be extant. Neither has the author been able to find the example mentioned in the literature, or to discover its exact nature. Consequently, some of the results obtainable may be of sufficient interest to warrant their exposition.

2. **Analytic functions with property** (A). Let there be given any set E of rational points in the complex plane; that is, a set of points of the form $a + bi$, when a and b are rational. Let

$$z_1, z_2, \cdots$$

be an arrangement in denumerable order of the points of the set. Let z_0 be any other point of the complex plane. We prove the following lemma.

Lemma. *There exists an entire function of z assuming rational values at points of E and assuming a transcendental value at z_0.*

Let

$$P_1(z) = (z - z_1)/N_1,$$

Received November 7, 1938.

where N_1 is an integer sufficiently large that $|P_1(z)| \leqq 1$ in $|z| \leqq 1$. Let

$$P_n(z) = P_{n-1}(z)\,(z - z_n)/N_n,$$

where N_n is an integer large enough that $|P_n(z)| \leqq 1$ in $|z| \leqq n$. Let $A_1, A_2, \cdots$ be non-zero rational complex numbers such that $\sum_1^\infty |A_n| < +\infty$. Then it is readily seen that

$$P(z) = \sum_1^\infty A_n P_n(z)$$

is an analytic function in the entire plane, for the series on the right converges uniformly in any bounded portion of the plane. Furthermore, $P(z)$ assumes rational complex values at all points of E. In fact,

$$P(z_n) = \sum_{k=1}^{n-1} A_k P_k(z_n) \qquad (n = 1, 2, 3, \cdots),$$

since $P_k(z_n) = 0$ for $k \geqq n$. Now $P_k(z_n)$ is rational, and so is A_k; hence $P(z_n)$ is rational.

Consider now

$$P(z_0) = \sum_{k=1}^\infty A_k P_k(z_0).$$

Since the A_k are at our disposal, we can always choose them so that

$$0 \leqq \text{argument } A_k P_k(z_0) < \frac{\pi}{2}.$$

That is, if $A_k P_k(z_0) = a_k + ib_k$, $a_k > 0$, $b_k \geqq 0$. Let

$$P(z_0) = \sum_{k=1}^\infty (a_k + ib_k) = a + ib.$$

If γ is any real transcendental number less than a, we can find real positive rational numbers c_k, each less than one, such that

$$\sum_1^\infty c_k a_k = \gamma.$$

In fact, put $\sum_{k=s+1}^\infty a_k = R_s$. Let $\sum_1^{n_1} a_k$ be $> \gamma$. Choose a positive rational number e_1, less than one, such that

$$0 < \gamma - e_1 \cdot \sum_1^{n_1} a_k < R_{n_1}.$$

Then let n_2 be such that

$$\gamma < e_1 \cdot \sum_1^{n_1} a_k + \sum_{n_1+1}^{n_2} a_k$$

and choose e_2 positive, rational, less than one, and such that

$$0 < \gamma - e_1 \cdot \sum_1^{n_1} a_k - e_2 \cdot \sum_{n_1+1}^{n_2} a_k < R_{n_2}.$$

Continuing in this way, we get rational numbers e_r, $0 < e_r < 1$, such that

$$0 < \gamma - e_1 \cdot \sum_1^{n_1} a_k - e_2 \cdot \sum_{n_1+1}^{n_2} a_k - \cdots - e_r \cdot \sum_{n_{r-1}+1}^{n_r} a_k < R_{n_r}.$$

It is obvious then that

$$e_1 \cdot \sum_1^{n_1} a_k + e_2 \cdot \sum_{n_1+1}^{n_2} a_k + e_3 \cdot \sum_{n_2+1}^{n_3} a_k + \cdots = \gamma;$$

or

$$\sum_1^{\infty} c_k a_k = \gamma,$$

where $c_k = e_i$ for $n_{i-1} < k \leqq n_i$; $n_0 = 0$. Thus

$$\sum_1^{\infty} A_k c_k P_k(z_0) = \sum_1^{\infty} c_k(a_k + ib_k) = \gamma + ib'.$$

Let then

$$\overline{P}(z) = \sum_1^{\infty} (A_k c_k) P_k(z).$$

$\overline{P}(z)$ will possess the same properties that were mentioned for $P(z)$, since it is obtained by replacing A_k in $P(z)$ by $A_k c_k$. Furthermore $\overline{P}(z_0) = \gamma + ib'$, which is trancendental. This completes the proof of the lemma.

THEOREM 1. *There exist entire functions which possess the property* (A) *with respect to all rational complex numbers, and which are not rational functions.*

In the lemma, take for E the set of all rational complex numbers, and for z_0 take an irrational algebraic number, as $\sqrt{2}$. Then the $\overline{P}(z)$ obtained will not be rational, for if it were, it would have to have rational coefficients and so would not assume a transcendental value at $z = \sqrt{2}$.

It may be remarked that if $P_n(z)$ is defined in a slightly different way, namely, as

$$P_n(z) = P_{n-1}(z)(z - z_n)(z - \bar{z}_n)/N_n,$$

where $\bar{z}$ is $|\ z\ |^2/z$, the functions $P_n(z)$ are all polynomials with real rational coefficients, and if the A_n are chosen real, $P(z)$ will be real for real z.

THEOREM 2. *Let E_1 and E_2 be any mutually disjoint sets of rational points. Then there exist analytic functions assuming rational values at points of E_1 and assuming rational values at no point of E_2.*

Let

$$E_1 = z_1, z_2, z_3, \cdots$$

and

$$E_2 = \zeta_1, \zeta_2, \zeta_3, \cdots.$$

By the lemma, there exists a function $P_i(z)$ which assumes rational values on $E_1 + (E_2 - \zeta_i)$ and which assumes at ζ_i an irrational value α_i. Let

$$Q_0(z) = P_1(z)/N_1,$$

and

$$Q_n(z) = P_{n+1}(z) \cdot \prod_{k=1}^{n} (z - z_k) \cdot (z - \zeta_k)/N_n,$$

where N_n is an integer large enough that $| Q_n(z) | \leqq 1$ in $| z | \leqq n$. Then

$$Q(z) = \sum_{n=0}^{\infty} \frac{1}{2^n} Q_n(z)$$

possesses the properties stated in the theorem. In fact,

$$Q(z_n) = \sum_{k=0}^{n-1} \frac{1}{2^k} Q_k(z_n),$$

and so $Q(z_n)$ is rational. Similarly,

$$Q(\zeta_n) = \sum_{k=0}^{n-1} \frac{1}{2^k} Q_k(\zeta_n) = R + \frac{\alpha_n}{2^{n-1} N_{n-1}} \prod_{k=1}^{n-1} (\zeta_n - z_k)(\zeta_n - \zeta_k),$$

where R represents a rational number. Hence $Q(\zeta_n)$ is irrational.

In connection with the functions with property (A), we may ask the following questions. Let the function

$$P(z) = a_0 + a_1 z + a_2 z^2 + \cdots$$

possess the property (A) with respect to all rational complex numbers; the series is supposed to converge for all z. Then is it possible that

(1) $a_0, a_1, \cdots$ be rational without $P(z)$ being a rational function?

(2) some of the coefficients in the series be irrational?

THEOREM 3. *The questions* (1) *and* (2) *above may be answered affirmatively.*

To show that the a_k may be rational without $P(z)$ being a rational function, let us construct a function as follows: If $z_1, z_2, \cdots$ represents a denumerable arrangement of all the rational complex points, put

$$P_1(z) = (z - z_1)/N_1,$$

$$P_n(z) = P_{n-1}(z)(z - z_n) z^{n-1}/N_n,$$

where N_n is an integer large enough that $| P_n(z) | \leqq 1$ in $| z | \leqq n$. Let

$$P(z) = \sum_{1}^{\infty} A_n P_n(z).$$

The A_n are chosen as in the lemma to insure convergence of the series to a non-rational function with the property (A). Now

$$P^{(k)}(z) = \sum_{n=1}^{\infty} A_n P_n^{(k)}(z).$$

From the formula for $P_n(z)$, we see that $P_n^{(k)}(0) = 0$, for $n > k + 1$. Hence

$$P^{(k)}(0) = \sum_{n=1}^{k+1} A_n P_n^{(k)}(0),$$

or $P_k(0)$ is rational. Hence the coefficients of the power series for $P(z)$ will all be rational.

To show that it is possible that the power series expansion of a function with property (A) have some irrational coefficients, we perform the construction of a function $P(z)$ as in the lemma and Theorem 1 except that instead of choosing the A_k such that $\sum_1^{\infty} A_k P_k(z_0)$ is transcendental, we choose them so that $\sum_1^{\infty} A_k P_k'(0)$ is irrational. Then $P(z)$ will have the second coefficient in its power series irrational.

3. **Real functions of a real variable with property** (A). As was stated in the introduction, the familiar examples of continuous functions with property (A) are either built up piecewise from analytic ones, or are extremely irregular with respect to differentiability. In this section we shall show that, given k, a non-negative integer, there exists a function with property (A) on the interval (0, 1), with a continuous derivative of order k and with a derivative of order $k + 1$ at no point. That is, we shall show that functions with property (A) can be as regular or irregular as we desire.

There are several ways to construct such functions. The simplest, in principle at least, would be to construct a continuous function without a derivative, and attempt to show that it and its integrals possess the property (A). We shall adopt this procedure.

We define a function $f_n(x)$ as follows. Let S_n be the set of points between zero and one of the form $m/(2^{n-1}n!)$ $(m = 0, 1, \cdots, 2^{n-1}n!)$. Let $p < q$ be two consecutive points of S_n; then $f_n(x)$ is defined for $p \leqq x \leqq q$ as the continuous function vanishing at p and q, and with derivative $+1$ for $p < x < \frac{1}{2}(p + q)$ and derivative -1 for $\frac{1}{2}(p + q) < x < q$. Then $0 \leqq f_n(x) \leqq (2^n n!)^{-1}$, and $\sum_1^{\infty} f_n(x)$ converges uniformly. We define

$$f(x) = \sum_{n=1}^{\infty} f_n(x).$$

It follows from well-known principles that $f(x)$ possesses a derivative at no point.[1] We consider now the integrals of $f(x)$, and those of $f_n(x)$, since the series may be integrated termwise.

[1] See, for example, Hobson, *The Theory of Functions of a Real Variable*, 1907, p. 621.

To obtain a general result, let $g(x)$ be a function defined in $0 \leqq x \leqq 1$, periodic with period $i = 1/r$, r being a positive integer. Let $g_1(x)$, $g_2(x)$, $\cdots$ be the successive integrals of $g(x)$. We see easily that

$$g_1(x) = h_1(x) + S_1(x),$$

where $h_1(x) = \int_0^x g(t)dt$ for $0 \leqq x \leqq i$ and $h_1(x + i) = h_1(x)$ for $x > 0$, and where $S_1(x)$ is a step function defined as 0 for $0 \leqq x \leqq i$ and $S_1(x + i) = S_1(x) + a_1$ for $x > 0$, where $a_1 = \int_0^i g(t)dt$. Now $h_1(x)$ is another periodic function such as $g(x)$ (except at $x = 0$, and the value of $h_1(x)$ at this point is of no importance), and so we see that

$$g_2(x) = h_2(x) + S_2(x) + S_1^1(x),$$

where

(a) $h_2(x) = \int_0^x h_1(t)dt$ for $0 \leqq x \leqq i$ and $h_2(x + i) = h_2(x)$ for $x > 0$;

(b) $S_2(x) = 0$ for $0 \leqq x \leqq i$ and $S_2(x + i) = S_2(x) + a_2$ for $x > 0$,

where $a_2 = \int_0^i h_1(t)dt$;

(c) $S_1^1(x) = \int_0^x S_1(t)\,dt$.

By continuation of this process, we get

$$g_k(x) = h_k(x) + S_k(x) + S_{k-1}^1(x) + \cdots + S_1^{k-1}(x), \tag{1}$$

where

(a) $h_k(x) = \int_0^x h_{k-1}(t)\,dt$ for $0 \leqq x \leqq i$ and $h_k(x + i) = h_k(x)$ for $x > 0$;

(b) $S_k(x) = 0$ for $0 \leqq x \leqq i$ and $S_k(x + i) = S_k(x) + a_k$, $x > 0$;

(2)
$$a_k = \int_0^i h_{k-1}(t)\,dt;$$

(c) $S_i^j(x) = \int_0^x S_i^{j-1}(t)\,dt$.

We now find a formula for the successive integrals of a step function.

LEMMA. *If $S(x)$ is a step function defined as $S(x) = 0$ $(0 \leqq x \leqq i)$ and $S(x + i) = S(x) + a$ for $x > 0$, and if*

$$S^1(x) = \int_0^x S(t)\,dt, \quad \cdots, \quad S^r(x) = \int_0^x S^{r-1}(t)\,dt,$$

then

$$S^r(mi) = \frac{ai^r}{r!}\sum_{j=1}^{m-1} j^r. \tag{3}$$

We have

$$S^1(x) = \int_0^x S(t)\,dt,$$

$$S^2(x) = \int_0^x dt \int_0^t S(\alpha)\,d\alpha = \int_0^x S(t)(x-t)\cdot dt,$$

$$S^r(x) = \int_0^x dt \int_0^t S^{r-1}(\alpha)\,d\alpha = \int_0^x S(t)\,\frac{(x-t)^{r-1}}{(r-1)!}\,dt.$$

Integrating this last by parts, we obtain

$$S^r(x) = -S(t)\,\frac{(x-t)^r}{r!}\Big]_0^x + \int_0^x \frac{(x-t)^r}{r!}\,dS(t).$$

The integrated term vanishes, and if we put $x = mi$, we obtain immediately the stated result.

If we use (3), (1) reduces to

$$g_k(mi) = ma_k + \frac{a_{k-1}i}{1!}\sum_{j=1}^{m-1} j + \frac{a_{k-2}i^2}{2!}\sum_{j=1}^{m-1} j^2 + \cdots + \frac{a_1 i^{k-1}}{(k-1)!}\sum_{j=1}^{m-1} j^{k-1}, \tag{4}$$

$$a_r = \int_0^i \cdots \int_0^{\alpha_3}\int_0^{\alpha_2} g(\alpha_1)\,d\alpha_1 \cdots d\alpha_r.$$

We apply these results to $f_n(x)$. The intervals in this case are of length $i_n = (2^{n-1}n!)^{-1}$. By direct integration of $f_n(x)$, which equals x for $0 \leqq x \leqq \frac{1}{2}i_n$, and equals $i_n - x$ for $\frac{1}{2}i_n \leqq x \leqq i_n$, it is but a simple exercise to show that

$$a_r^n = \int_0^{i_n} \cdots \int_0^{\alpha_3}\int_0^{\alpha_2} f_n(\alpha_1)\,d\alpha_1 \cdots d\alpha_r = \frac{i_n^{r+1}}{2^{r+1}}\sum_{s=1}^{r}\frac{1}{s!(r-s+1)!} = i_n^{r+1}b_r.$$

Thus

$$f_n^k(mi_n) = mi_n^{k+1}b_k + i_n^{k+1}\frac{b_{k-1}}{1!}\sum_{j=1}^{m-1} j + i_n^{k+1}\frac{b_{k-2}}{2!}\sum_{j=1}^{m-1} j^2 + \cdots + i_n^{k+1}\frac{b_1}{(k-1)!}\sum_{j=1}^{m-1} j^{k-1}. \tag{5}$$

Consider now a rational number $x = p/q$. For $n \geqq q$, the set of points S_n contains x, and the number of intervals between points of S_n, preceding x, is $x\cdot 2^{n-1}\cdot n!$. In equation (5) we put $m = x\cdot 2^{n-1}n!$, $i_n = (2^{n-1}n!)^{-1}$ and we get the value of $f_n^k(x)$. Let us examine one term of the right member, say

$$\left(\frac{b_{k-t}}{t!}\right) i_n^{k+1}\sum_{j=1}^{x/i_n-1} j^t. \tag{6}$$

The summation in (6) is a polynomial of degree $t+1$ in x/i_n, hence when $b_{k-t}i_n^{t+1}/t!$ is multiplied in, there results a polynomial of degree $t+1$ in x with coefficients that are products of powers of i_n and numbers not dependent on n; the lowest power of i_n appearing as a factor in one of these coefficients is $k+1-(t+1)=k-t$. We see then that the expression on the right side of (5) is of the form

$$A_1(k,n)x + A_2(k,n)x^2 + \cdots + A_k(k,n)x^k = f_n^k(x),$$

where the A's are polynomials in i_n of degree at least one and with coefficients rational numbers depending only on k. Now

$$f^k(x) = \sum_{n=1}^{q-1} f_n^k(x) + \sum_{n=q}^{\infty} f_n^k(x)$$

$$= \sum_{n=1}^{q-1} f_n^k(x) + \sum_{n=q}^{\infty} (A_1(k,n)x + \cdots + A_k(k,n)x^k).$$

Put

$$\sum_{n=0}^{\infty} A_j(k,n) = I_j(k), \qquad \text{and} \qquad I_1(k)x + \cdots + I_k(k)x^k = P_k(x).$$

Then

$$f^k(x) = \sum_{n=1}^{q-1} f_n^k(x) - \sum_{n=0}^{q-1} (A_1(k,n)x + \cdots + A_k(k,n)x^k) + P_k(x).$$

We see then that

$$f^k(x) - P_k(x)$$

assumes rational values at all rational points. We may state the results in a theorem.

THEOREM 4. *A function possessing property* (A), *with continuous k-th derivative and with $(k+1)$-th derivative nowhere, can be obtained by integrating $f(x)$ k times and subtracting a polynomial of degree k, possibly with irrational coefficients.*

Example. $P_1(x) = \frac{1}{2}e^{\frac{1}{2}}x$; thus

$$\int_0^x f(t)\,dt - \tfrac{1}{2}e^{\frac{1}{2}}x$$

is a function with continuous first derivative $f(x) - \frac{1}{2}e^{\frac{1}{2}}$ and with a second derivative at no point.

HARVARD UNIVERSITY.

A REMARK ON THE NORMAL DECOMPOSITIONS OF GROUPS

By Oystein Ore

In a recent paper[1] I have considered the representations of a group G as the union of normal subgroups

$$G = [A_1, A_2, \cdots, A_n], \tag{1}$$

where the A_i are normally indecomposable in G; i.e., they are not the union of two proper subgroups which are normal in G.

The question arises whether the A_i may be normally decomposable in themselves, i.e., whether there exists a representation

$$A = [B_1, B_2, \cdots, B_l], \tag{2}$$

where the B_i are proper normal subgroups in A. We shall prove the following

THEOREM.[2] *A component A_i of non-Abelian type in a normal indecomposable representation* (1) *is also indecomposable in itself.*

To prove this theorem let

$$B_i^{(1)}, B_i^{(2)}, \cdots$$

be the various conjugates of a B_i occurring in some representation (2). All these conjugates are also normal in A, and their union is normal in G. Since A is normally indecomposable in G, this means that there exists one B in (2) such that

$$A = [B^{(1)}, B^{(2)}, \cdots]$$

is the union of indecomposable conjugate groups.[3]

Now let $N^{(i)}$ be the unique maximal normal subgroup of A contained in $B^{(i)}$. All $N^{(i)}$ are conjugate and the simple groups

$$L_i = B^{(i)}/N^{(i)}$$

are all isomorphic. Furthermore, the union

$$C_A = [N^{(1)}, N^{(2)}, \cdots]$$

Received November 25, 1938.

[1] O. Ore, *Structures and group theory*, II, this Journal, vol. 4(1938), pp. 247–269.

[2] In the paper mentioned above it was indicated (p. 260, lines 18–21) that such a theorem could be proved. In this statement the condition that A should be of non-Abelian type had, however, inadvertently been omitted.

[3] Of course not all conjugates of the indecomposable B need to appear in the reduced form of this representation, but we assume that the descending chain condition holds for the normal subgroups in A such that the reduced representation contains only a finite number of components.

is the upper normal cover group of A. Since C_A is a characteristic subgroup of A, it is normal in G.

Now let M be the unique maximal normal subgroup of G contained in A. Then M must contain C_A. The quotient group A/C_A is known to be the direct product of simple groups each isomorphic to L_i. This implies that also M/C_A is the product of simple groups. Furthermore, the product representation of M/C_A can be completed into a direct product representation for A/C_A

$$A/C_A = M/C_A \times K/C_A . \tag{3}$$

Here K is normal in A and the group K/C_A is simple and isomorphic to any L_i and to A/M.

In (3) the group K cannot be normal in G since it would make A normally decomposable in G. Therefore by transforming (3) with the various elements of G one obtains different basis for A/C_A. This is, however, only possible when the simple groups are cyclic of prime order, and we have assumed that this is not the case.

YALE UNIVERSITY.

ON IMBEDDING A SPACE IN A COMPLETE SPACE

By L. W. Cohen

If S is a metric space, a sequence $p_k \epsilon S$ is a Cauchy sequence if for every $\epsilon > 0$ there is a k_ϵ such that, for $k \geqq k_\epsilon$, $h > 0$, $\rho(p_k, p_{k+h}) < \epsilon$. If $S(p, \epsilon)$ is the sphere of radius ϵ, this condition may be expressed as

$$p_k \epsilon S(p_{k_\epsilon}, \epsilon), \qquad k \geqq k_\epsilon .$$

It is a classical result that for any metric S there exists a complete metric S^* such that S is homeomorphic to a subset of S^* and the image of a Cauchy sequence in S is a convergent sequence in S^*. If we consider a class S of elements p, called points, and a class A of elements α, called indices, such that for each $p \epsilon S$ and $\alpha \epsilon A$ there is a well defined subset $U_\alpha(p) \subset S$, we may study an analogous problem. A sequence $p_k \epsilon S$ may be called a Cauchy sequence if for every $\alpha \epsilon A$ there are a $q_\alpha \epsilon S$ and a $k_\alpha > 0$ such that

$$p_k \epsilon U_\alpha(q_\alpha), \qquad k \geqq k_\alpha .$$

The problem is that of formulating conditions upon the $U_\alpha(p)$ such that a theorem analogous to the one given above for metric S may be proved. We shall give a solution of this problem in this paper. The method is related to that used by Cantor in defining the real numbers as classes of equivalent sequences of rational numbers. The conditions are related to those of Chittenden, Alexandroff and Urysohn, Niemytski and Weil in their work on the metrization problem and uniformity properties of topological spaces.

The space S, the index set A and the sets $U_\alpha(p)$, called neighborhoods, are subjected to the following postulates:

I. $\prod_{\alpha \epsilon A} U_\alpha(p) = p$.

II. If $p \epsilon S$ and $\alpha, \beta \epsilon A$, there is $\gamma = \gamma(\alpha, \beta; p)$ such that $U_\alpha(p) U_\beta(p) \supset U_\gamma(p)$.

III. If $p \epsilon S$ and $\alpha \epsilon A$, then there are $\lambda(\alpha)$, $\delta(p, \alpha) \epsilon A$ such that, if $q \epsilon S$ and

$$U_{\delta(p,\alpha)}(q) U_{\lambda(\alpha)}(p) \neq 0,$$

then

$$U_{\delta(p,\alpha)}(q) \subset U_\alpha(p).$$

We shall refer to $\lambda(\alpha)$ as the first index of III and $\delta(p, \alpha)$ as the second index of III.

If $M \subset S$, $\bar{M}$ is the set of all $p \epsilon S$ such that, for all $\alpha \epsilon A$, $U_\alpha(p) M \neq 0$.

Received January 3, 1939.

THEOREM 1. *If $p \in S$, $\alpha \in A$ and $\lambda(\alpha)$ is the first index of* III, *then* $\overline{U}_{\lambda(\alpha)}(p) \subset U_\alpha(p)$.

Proof. If $q \in \overline{U}_{\lambda(\alpha)}(p)$, then for all $\beta \in A$, $U_\beta(q)U_{\lambda(\alpha)}(p) \neq 0$. In particular this is so for $\beta = \delta(p, \alpha)$, the second index of III. Thus from I and III

$$q \in U_{\delta(p,\alpha)}(q) \subset U_\alpha(p),$$

and so

$$\overline{U}_{\lambda(\alpha)}(p) \subset U_\alpha(p).$$

We shall study families $\{U_\alpha(p_\alpha)\}$ of neighborhoods containing just one $U_\alpha(p_\alpha)$ for each $\alpha \in A$ and such that, for every finite set $\{\alpha_1, \cdots, \alpha_n\} \subset A$,

$$\prod_{i=1}^{n} U_{\alpha_i}(p_{\alpha_i}) \neq 0. \tag{1}$$

Such a family will be denoted by $\Pi = \{U_\alpha(p_\alpha)\}$ and the product (1) by $\Pi(\alpha_1, \cdots, \alpha_n)$. It is clear that if $\{\alpha_1, \cdots, \alpha_n\} \supset \{\beta_1, \cdots, \beta_m\}$ then $\Pi(\alpha_1, \cdots, \alpha_n) \subset \Pi(\beta_1, \cdots, \beta_m)$. We will call two families $\Pi' = \{U_\alpha(p'_\alpha)\}$ and $\Pi'' = \{U_\alpha(p''_\alpha)\}$ equivalent and write $\Pi' \sim \Pi''$ if for every $\alpha \in A$ there are $p_\alpha \in S$ and a set $\{\alpha_1, \cdots, \alpha_n\} \subset A$ such that

$$\Pi'(\alpha_1, \cdots, \alpha_n) + \Pi''(\alpha_1, \cdots, \alpha_n) \subset U_\alpha(p_\alpha). \tag{2}$$

THEOREM 2. *If $\Pi' \sim \Pi''$, there is a Π such that for every $\{\alpha_1, \cdots, \alpha_n\} \subset A$ there is $\{\beta_1, \cdots, \beta_m\} \subset A$ such that*

$$\Pi'(\beta_1, \cdots, \beta_m) + \Pi''(\beta_1, \cdots, \beta_m) \subset \Pi(\alpha_1, \cdots, \alpha_n).$$

Proof. For any $\alpha_j \in A$ $(j = 1, \cdots, n)$, let $\alpha_{ij} = \alpha_i(\alpha_j)$ $(i = 1, \cdots, m_j)$ be the indices satisfying (2). Denoting the $k = n(m_1 + \cdots + m_n)$ indices α_{ij} by $\beta_1, \cdots, \beta_k$ we get

$$\Pi'(\beta_1, \cdots, \beta_k) + \Pi''(\beta_1, \cdots, \beta_k) \subset \prod_{j=1}^{n} U_{\alpha_j}(p_{\alpha_j}).$$

Thus the $U_\alpha(p_\alpha)$ of (2) define a Π satisfying the theorem.

THEOREM 3. *If Π' and Π'' are such that for each $\{\alpha_1, \cdots, \alpha_n\} \subset A$*

$$\Pi'(\alpha_1, \cdots, \alpha_n)\Pi''(\alpha_1, \cdots, \alpha_n) \neq 0,$$

then $\Pi' \sim \Pi''$.

Proof. Let $\Pi' = \{U_\alpha(p'_\alpha)\}$ and $\Pi'' = \{U_\alpha(p''_\alpha)\}$. For $\alpha \in A$ let $\lambda(\alpha)$ be the first index of III and $\delta(\alpha) = \delta(p'_{\lambda(\alpha)}, \alpha)$ be the second index of III. Now

$$0 \neq \Pi''(\delta(\alpha), \lambda(\alpha))\Pi'(\delta(\alpha), \lambda(\alpha)) \subset U_{\delta(\alpha)}(p''_{\delta(\alpha)})U_{\lambda(\alpha)}(p'_{\lambda(\alpha)}).$$

From III we have $U_{\delta(\alpha)}(p''_{\delta(\alpha)}) \subset U_\alpha(p'_{\lambda(\alpha)})$ and from Theorem 1 $U_{\lambda(\alpha)}(p'_{\lambda(\alpha)}) \subset U_\alpha(p'_{\lambda(\alpha)})$. Hence

$$\Pi''(\delta(\alpha), \lambda(\alpha)) + \Pi'(\delta(\alpha), \lambda(\alpha)) \subset U_{\delta(\alpha)}(p''_{\delta(\alpha)}) + U_{\lambda(\alpha)}(p'_{\lambda(\alpha)}) \subset U_\alpha(p'_{\lambda(\alpha)})$$

and $\Pi'' \sim \Pi'$.

THEOREM 4. *If* $\Pi = \{U_\alpha(p_\alpha)\}$ *and, for each* α, $\lambda(\alpha)$ *is the first index of* III, *then* $\{U_\alpha(p_{\lambda(\alpha)})\}$ *defines a* Π^* *such that* $\Pi \sim \Pi^*$; *and if* $\Pi' \sim \Pi$, $\{\alpha_1, \cdots, \alpha_n\} \subset A$, *then there are* $\beta_i = \beta_i(\alpha_1, \cdots, \alpha_n ; \Pi')$ $(i = 1, \cdots, m)$ *such that* $\Pi'(\beta_1, \cdots, \beta_m) \subset \Pi^*(\alpha_1, \cdots, \alpha_n)$.

Proof. For any $\alpha \in A$, $U_{\lambda(\alpha)}(p_{\lambda(\alpha)}) \subset U_\alpha(p_{\lambda(\alpha)})$ follows from Theorem 1, and for every $\{\alpha_1, \cdots, \alpha_n\} \subset A$,

$$0 \neq \Pi(\lambda(\alpha_1), \cdots, \lambda(\alpha_n)) \subset \prod_{i=1}^{n} U_{\alpha_i}(p_{\lambda(\alpha_i)}).$$

Thus the $U_\alpha(p_{\lambda(\alpha)})$ define a family we may denote by Π^*. If $\{\alpha_1, \cdots, \alpha_n\} \subset A$, we consider $\lambda_i = \lambda(\alpha_i)$, p_{λ_i}, $\delta_i = \delta(p_{\lambda_i}, \alpha_i)$ for $i = 1, \cdots, n$, where λ_i and δ_i are the first and second indices of III associated with α_i and p_{λ_i}. Then since $\Pi \sim \Pi'$, there are a $\Pi'' = \{U_\alpha(q_\alpha)\}$ and $\{\beta_1, \cdots, \beta_k\} \subset A$ such that

$$\Pi(\beta_1, \cdots, \beta_k) + \Pi'(\beta_1, \cdots, \beta_k) \subset \Pi''(\delta_1, \cdots, \delta_n) = \prod_{i=1}^{n} U_{\delta_i}(q_{\delta_i}) \tag{3}$$

according to Theorem 2. Now

$$\Pi(\beta_1, \cdots, \beta_k, \lambda_1, \cdots, \lambda_n) \subset \Pi(\beta_1, \cdots, \beta_k) \subset U_{\delta_i}(q_{\delta_i}),$$

$$\Pi(\beta_1, \cdots, \beta_k, \lambda_1, \cdots, \lambda_n) \subset U_{\lambda_i}(p_{\lambda_i}), \qquad (i = 1, 2, \cdots, n).$$

Hence $U_{\delta_i}(q_{\delta_i})U_{\lambda_i}(p_{\lambda_i}) \neq 0$ and $U_{\delta_i}(q_{\delta_i}) \subset U_{\alpha_i}(p_{\lambda(\alpha_i)})$, since $\delta_i = \delta(p_{\lambda(\alpha_i)}, \alpha_i)$, $\lambda_i = \lambda(\alpha_i)$. Thus from (3)

$$\Pi'(\beta_1, \cdots, \beta_k) \subset \prod_{i=1}^{n} U_{\delta_i}(q_{\delta_i}) \subset \prod_{i=1}^{n} U_{\alpha_i}(p_{\lambda(\alpha_i)}) = \Pi^*(\alpha_1, \cdots, \alpha_n).$$

Clearly

$$0 \neq \Pi(\lambda(\alpha_1), \cdots, \lambda(\alpha_n), \alpha_1, \cdots, \alpha_n) \subset \Pi(\alpha_1, \cdots, \alpha_n)\Pi^*(\alpha_1, \cdots, \alpha_n),$$

and, from Theorem 3, $\Pi \sim \Pi^*$.

THEOREM 5. *If* $p \in S$, *the family* $\{U_\alpha(p)\}$ *defines a* $\Pi = \Pi^p$ *which satisfies Theorem* 4 *with respect to* $\Pi' \sim \Pi^p$.

Proof. From I it follows that $p \in \Pi^p(\alpha_1, \cdots, \alpha_n)$ for all $\{\alpha_1, \cdots, \alpha_n\} \subset A$. Since for all α, $p = p_\alpha = p_{\lambda(\alpha)}$, $\Pi^p = \Pi^*$ in Theorem 4.

THEOREM 6. *The relation* $\Pi' \sim \Pi''$ *is reflexive, symmetric and transitive.*

Proof. The first two properties of the equivalence relation are clear. Suppose that $\Pi^{(1)} \sim \Pi^{(2)}$ and $\Pi^{(2)} \sim \Pi^{(3)}$. There are Π', Π'' such that, by Theorem 2, for every $\{\alpha_1, \cdots, \alpha_n\} \subset A$, there are $\{\beta_1, \cdots, \beta_m\}$, $\{\gamma_1, \cdots, \gamma_k\} \subset A$ such that

$$\begin{aligned} &\Pi^{(1)}(\beta_1, \cdots, \beta_m) + \Pi^{(2)}(\beta_1, \cdots, \beta_m) \subset \Pi'(\alpha_1, \cdots, \alpha_n), \\ &\Pi^{(2)}(\gamma_1, \cdots, \gamma_k) + \Pi^{(2)}(\gamma_1, \cdots, \gamma_k) \subset \Pi''(\alpha_1, \cdots, \alpha_n). \end{aligned} \tag{4}$$

Hence

$$0 \neq \Pi^{(2)}(\beta_1, \cdots, \beta_m, \gamma_1, \cdots, \gamma_k) \subset \Pi'(\alpha_1, \cdots, \alpha_n)\Pi''(\alpha_1, \cdots, \alpha_n),$$

and $\Pi' \sim \Pi''$ by Theorem 3. For any $\alpha \in A$ there are then $p_\alpha \in S$ and $\alpha_i(\alpha) = \alpha_i \in A$ $(i = 1, \cdots, n)$ such that

$$\Pi'(\alpha_1, \cdots, \alpha_n) + \Pi''(\alpha_1, \cdots, \alpha_n) \subset U_\alpha(p_\alpha). \tag{5}$$

From (4) and (5) there are $\beta_1, \cdots, \beta_m, \gamma_1, \cdots, \gamma_k \in A$ such that

$$\Pi^{(1)}(\beta_1, \cdots, \beta_m, \gamma_1, \cdots, \gamma_k) + \Pi^{(3)}(\beta_1, \cdots, \beta_m, \gamma_1, \cdots, \gamma_k) \subset U_\alpha(p_\alpha)$$

and $\Pi^{(1)} \sim \Pi^{(3)}$.

On the basis of Theorem 6 all families $\Pi = \{U_\alpha(p_\alpha)\}$ fall into mutually exclusive classes $C(\Pi)$ such that $\Pi' \in C(\Pi)$ if and only if $\Pi' \sim \Pi$. To each $p \in S$ there is a class $C(\Pi^p)$, where $\Pi^p = \{U_\alpha(p)\}$. Let S^* be the class of all points $P = C(\Pi)$. From each $P \in S^*$ we choose a Π^P satisfying the conclusion of Theorem 4. As neighborhoods of $P \in S^*$ we define sets $U^*_{\alpha_1 \cdots \alpha_n}(P)$ for each $\{\alpha_1, \cdots, \alpha_n\} \subset A$ consisting of all $Q \in S^*$ such that for some $\{\beta_1, \cdots, \beta_m\} \subset A$

$$\Pi^Q(\beta_1, \cdots, \beta_m) \subset \Pi^P(\alpha_1, \cdots, \alpha_n).$$

THEOREM 7. *The space S^* is a regular Hausdorff space with the property:*

III*. *For each $\{\alpha_1, \cdots, \alpha_n\} \subset A$ there is $\{\gamma_1, \cdots, \gamma_k\} \subset A$ such that if*

$$U^*_{\gamma_1 \cdots \gamma_k}(Q)U^*_{\gamma_1 \cdots \gamma_k}(P) \neq 0,$$

then

$$U^*_{\gamma_1 \cdots \gamma_k}(Q) \subset U^*_{\alpha_1 \cdots \alpha_n}(P).$$

Proof. From the definition of the neighborhoods of P it is clear that

$$P \in \prod_{\{\alpha_1 \cdots \alpha_n\} \subset A} U^*_{\alpha_1 \cdots \alpha_n}(P).$$

Suppose that for every $\{\alpha_1, \cdots, \alpha_n\} \subset A$, $U^*_{\alpha_1 \cdots \alpha_n}(P)U^*_{\alpha_1 \cdots \alpha_n}(Q) \neq 0$. Then for every $\{\alpha_1, \cdots, \alpha_n\} \subset A$

$$\Pi^Q(\alpha_1, \cdots, \alpha_n)\Pi^P(\alpha_1, \cdots, \alpha_n) \neq 0,$$

and by Theorem 2, $\Pi^Q = \Pi^P$, $C(\Pi^Q) = C(\Pi^P)$ and $Q = P$. Hence for $Q \neq P$ there are disjoint neighborhoods of P and Q. Since

$$\Pi^P(\alpha_1, \cdots, \alpha_n)\Pi^P(\beta_1, \cdots, \beta_m) = \Pi^P(\alpha_1, \cdots, \alpha_n, \beta_1, \cdots, \beta_m),$$

$$U^*_{\alpha_1 \cdots \alpha_n}(P)U^*_{\beta_1 \cdots \beta_m}(P) = U^*_{\alpha_1 \cdots \alpha_n \beta_1 \cdots \beta_m}(P).$$

If $Q \in U^*_{\alpha_1 \cdots \alpha_n}(P)$, then, for some $\{\beta_1, \cdots, \beta_m\} \subset A$, $\Pi^Q(\beta_1, \cdots, \beta_m) \subset \Pi^P(\alpha_1, \cdots, \alpha_n)$ and $U^*_{\beta_1 \cdots \beta_m}(Q) \subset U^*_{\alpha_1 \cdots \alpha_n}(P)$. Thus S^* is a Hausdorff space.

To establish III* consider $P \in S^*$, $\{\alpha_1, \cdots, \alpha_n\} \subset A$, $\Pi^P = \{U_\alpha(p_\alpha)\}$ and,

for α_i, the first index $\lambda_i = \lambda(\alpha_i)$ of III. For α_i and λ_i consider p_{λ_i} and the second index $\delta_i = \delta(p_{\lambda_i}, \alpha_i)$ of III. From Theorem 1 we have

$$0 \neq \Pi^P(\lambda_1, \cdots, \lambda_n) = \prod_{i=1}^{n} U_{\lambda_i}(p_{\lambda_i}) \subset \prod_{i=1}^{n} U_{\alpha_i}(p_{\lambda_i}).$$

Hence the family $\{U_\alpha(p_{\lambda(\alpha)})\}$ defines a $\tilde{\Pi}$. Since for $\{\alpha_1, \cdots, \alpha_n\} \subset A$

$$0 \neq \Pi^P\{\lambda_1, \cdots, \lambda_n, \alpha_1, \cdots, \alpha_n\} \subset \Pi^P(\alpha_1, \cdots, \alpha_n)\tilde{\Pi}(\alpha_1, \cdots, \alpha_n),$$

$\tilde{\Pi} \sim \Pi^P$ follows from Theorem 2. Thus from Theorem 4, for each $\{\alpha_1, \cdots, \alpha_n\} \subset A$ there is $\{\beta_1, \cdots, \beta_m\} \subset A$ such that

$$\tilde{\Pi}(\beta_1, \cdots, \beta_m) \subset \Pi^P(\alpha_1, \cdots, \alpha_n). \tag{6}$$

Now for each β_i let $\lambda_i = \lambda(\beta_i)$ be the first index of III, let p_{λ_i} be the $p_{\lambda(\beta_i)}$ of $U_{\lambda(\beta_i)}(p_{\lambda(\beta_i)})$ and let $\delta_i = \delta(p_{\lambda_i}, \beta_i)$ be the second index of III. Suppose that

$$U^*_{\gamma_1 \cdots \gamma_{2m}}(Q)U^*_{\gamma_1 \cdots \gamma_{2m}}(P) \neq 0 \qquad (\gamma_i = \lambda_i, \gamma_{m+i} = \delta_i; i = 1, \cdots, m). \tag{7}$$

There is $R \in U^*_{\gamma_1 \cdots \gamma_{2m}}(Q)U^*_{\gamma_1 \cdots \gamma_{2m}}(P)$ and so for some $\{\mu_1, \cdots, \mu_k\} \subset A$

$$0 \neq \Pi^R(\mu_1, \cdots, \mu_k) \subset \Pi^Q(\gamma_1, \cdots, \gamma_{2m})\Pi^P(\gamma_1, \cdots, \gamma_{2m})$$

$$\subset \Pi^Q(\delta_1, \cdots, \delta_m)\ \Pi^P(\lambda_1, \cdots, \lambda_m) = \prod_{i=1}^{m} U_{\delta_i}(q_{\delta_i}) \prod_{i=1}^{m} U_{\lambda_i}(p_{\lambda_i}),$$

where $\Pi^Q = \{U_\alpha(q_\alpha)\}$. From the definitions of λ_i, δ_i we have

$$U_{\delta_i}(q_{\delta_i})U_{\lambda_i}(p_{\lambda_i}) = U_{\delta(p_{\lambda(\beta_i)}, \beta_i)}(q_{\delta_i})U_{\lambda(\beta_i)}(p_{\lambda(\beta_i)}) \neq 0.$$

Hence by III

$$U_{\delta_i}(q_{\delta_i}) \subset U_{\beta_i}(p_{\lambda_i}) \qquad (i = 1, \cdots, m).$$

Thus

$$\Pi^Q(\delta_1, \cdots, \delta_m) = \prod_{i=1}^{m} U_{\delta_i}(q_{\delta_i}) \subset \prod_{i=1}^{m} U_{\beta_i}(p_{\lambda_i}) = \tilde{\Pi}(\beta_1, \cdots, \beta_m).$$

Combining this with (6) and the fact that $\{\gamma_1, \cdots, \gamma_{2m}\} \supset \{\delta_1, \cdots, \delta_m\}$, we have

$$\Pi^Q(\gamma_1, \cdots, \gamma_{2m}) \subset \Pi^Q(\delta_1, \cdots, \delta_m) \subset \tilde{\Pi}(\beta_1, \cdots, \beta_m) \subset \Pi^P(\alpha_1, \cdots, \alpha_n)$$

from which it follows that

$$U^*_{\gamma_1 \cdots \gamma_{2m}}(Q) \subset U^*_{\alpha_1 \cdots \gamma_n}(P). \tag{8}$$

Thus (7) implies (8) and III* is proved.

The regularity of S^* is a consequence of III*. For any $U^*_{\alpha_1 \cdots \alpha_n}(P)$ let $\{\gamma_1, \cdots, \gamma_k\} \subset A$ satisfy III*. If $Q \in \overline{U}^*_{\gamma_1 \cdots \gamma_k}(P)$, then $U^*_{\gamma_1 \cdots \gamma_k}(Q)U^*_{\gamma_1 \cdots \gamma_k}(P) \neq 0$ and

$$Q \in U^*_{\gamma_1 \cdots \gamma_k}(Q) \subset U^*_{\alpha_1 \cdots \alpha_n}(P).$$

A sequence $P_k \in S^*$ is a Cauchy sequence if for every $\{\alpha_1, \cdots, \alpha_n\} \subset A$ there are $Q_{\alpha_1 \cdots \alpha_n} \in S^*$ and a $k(\alpha_1, \cdots, \alpha_n) > 0$ such that

$$P_k \in U^*_{\alpha_1 \cdots \alpha_n}(Q_{\alpha_1 \cdots \alpha_n}), \qquad k \geqq k(\alpha_1, \cdots, \alpha_n).$$

We write $\lim_k P_k = P$ if

$$P_k \in U^*_{\alpha_1 \cdots \alpha_n}(P), \qquad k \geqq k(\alpha_1, \cdots, \alpha_n).$$

THEOREM 8. *The space S^* is complete.*

Proof. If P_k is a Cauchy sequence in S^*, then for each $\alpha \in A$ there are $Q_\alpha \in S^*$ and $k(\alpha) > 0$ such that

$$P_k \in U^*_\alpha(Q_\alpha), \qquad k \geqq k(\alpha).$$

Let $\Pi^{P_k} = \{U_\beta(p^k_\beta)\}$, $\Pi^{Q_\alpha} = \{U_\beta(q^\alpha_\beta)\}$. Then for each $k \geqq k(\alpha)$ there are $\beta^k_i = \beta^k_i(\alpha) \in A$ $(i = 1, \cdots, n(k, \alpha))$ such that

$$\Pi^{P_k}(\beta^k_1, \cdots, \beta^k_{n(k,\alpha)}) \subset \Pi^{Q_k}(\alpha) = U_\alpha(q^\alpha_\alpha).$$

Consider any $\{\alpha_1, \cdots, \alpha_m\} \subset A$, $k(\alpha_1, \cdots, \alpha_m) = \max_{1 \leq j \leq m} k(\alpha_j)$, $k \geqq k(\alpha_1, \cdots, \alpha_m)$. Let $\beta^k_{ij} = \beta^k_i(\alpha_j)$ $(i = 1, \cdots, n(k, \alpha_j); j = 1, \cdots, m)$; $N = n(k, \alpha_1) + \cdots + n(k, \alpha_m)$, and put

$$\{\beta^k_{1,1}, \cdots, \beta^k_{n(k,\alpha_1),1}, \cdots, \beta^k_{1,m}, \cdots, \beta^k_{n(k,\alpha_m),m}\} = \{\gamma^k_1, \cdots, \gamma^k_N\}.$$

Then

$$\Pi^{P_k}(\gamma^k_1, \cdots, \gamma^k_N) \subset \Pi^{P_k}(\beta^k_1(\alpha_j), \cdots, \beta^k_{n(k,\alpha_j)}(\alpha_j)) \subset U_{\alpha_j}(q^{\alpha_j}_{\alpha_j}) \qquad (j = 1, \cdots, m),$$

$$(9) \qquad 0 \neq \Pi^{P_k}(\gamma^k_1, \cdots, \gamma^k_N) \subset \prod_{j=1}^{m} U_{\alpha_j}(q^{\alpha_j}_{\alpha_j}), \qquad k \geqq k(\alpha_1, \cdots, \alpha_m).$$

Hence the $U_\alpha(q^\alpha_\alpha)$ define a Π and a $C(\Pi) = P \in S^*$. Let Π^P be the family of $C(\Pi)$ in terms of which the neighborhoods of P are defined. Then from Theorem 4, for every $\{\delta_1, \cdots, \delta_t\} \subset A$ there is $\{\alpha_1, \cdots, \alpha_m\} \subset A$ such that

$$(10) \qquad \Pi(\alpha_1, \cdots, \alpha_m) = \prod_{i=1}^{m} U_{\alpha_j}(q^{\alpha_j}_{\alpha_j}) \subset \Pi^P(\delta_1, \cdots, \delta_t).$$

For these α_j we may determine $k_0 = k(\alpha_1, \cdots, \alpha_m)$ such that (9) is satisfied for $k \geqq k_0$ and some $\{\gamma^k_1, \cdots, \gamma^k_N\} \subset A$. From (9) and (10)

$$\Pi^{P_k}(\gamma^k_1, \cdots, \gamma^k_N) \subset \Pi^P(\delta_1, \cdots, \delta_t), \qquad k \geqq k_0,$$

$$P_k \in U^*_{\delta_1 \cdots \delta_t}(P), \qquad \lim_k P_k = P.$$

The neighborhoods for the points of S^* are determined by the choice of Π^* from $P = C(\Pi)$, the only requirement being that Π^* satisfies Theorem 4. It is easily seen that such are not unique. Let $\bar{S}$ be the space whose points are

the $C(\Pi)$, but whose neighborhoods are defined by $\tilde{\Pi}^P \in P$. The space S has the properties of Theorems 7 and 8.

THEOREM 9. *The spaces S^* and $\tilde{S}$ are homeomorphic.*

Proof. Let the neighborhoods in S^* be $U^*_{\alpha_1 \cdots \alpha_n}(P)$ and those in $\tilde{S}$ be $\tilde{U}_{\alpha_1 \cdots \alpha_n}(P)$. Since Π^P and $\tilde{\Pi}^P$ both satisfy Theorem 4, there are, for any $\{\alpha_1, \cdots, \alpha_n\} \subset A$, $\{\beta'_1, \cdots, \beta'_{m'}\} \subset A$ and $\{\beta''_1, \cdots, \beta''_{m''}\} \subset A$ such that

$$\tilde{\Pi}^P\{\beta'_1, \cdots, \beta'_{m'}\} \subset \Pi^P\{\alpha_1, \cdots, \alpha_n\},$$

$$\Pi^P\{\beta''_1, \cdots, \beta''_{m''}\} \subset \tilde{\Pi}^P\{\alpha_1, \cdots, \alpha_n\}.$$

From this it follows that for any $U^*_{\alpha_1 \cdots \alpha_n}(P)$ there is $\tilde{U}_{\beta'_1 \cdots \beta'_{m'}}(P) \subset U^*_{\alpha_1 \cdots \alpha_n}(P)$ and for any $\tilde{U}_{\alpha_1 \cdots \alpha_n}(P)$ there is $U^*_{\beta''_1 \cdots \beta''_{m''}}(P) \subset \tilde{U}_{\alpha_1 \cdots \alpha_n}(P)$. Thus S^* and $\tilde{S}$ are homeomorphic.

Let S^* be that one of the various homeomorphic complete spaces constructed above for which the neighborhoods of $P = C(\Pi^P)$ are determined by $\Pi^p = \{U_\alpha(p)\}$, $p \in S$. This is legitimate because of Theorem 5. We prove the following imbedding theorem.

THEOREM 10. *The mapping $f(p) = P = C(\Pi^p)$ on S to $f(S) \subset S^*$ is a homeomorphism such that if p_k is a Cauchy sequence in S, then $f(p_k)$ is a convergent sequence in S^*.*

Proof. If $C(\Pi^{p_1}) = C(\Pi^{p_2})$ then $\Pi^{p_1} \sim \Pi^{p_2}$ and for each $\{\alpha_1, \cdots, \alpha_n\} \subset A$ there is $\{\beta_1, \cdots, \beta_m\} \subset A$ such that

$$p_2 \in \Pi^{p_2}(\beta_1, \cdots, \beta_m) \subset \Pi^{p_1}(\alpha_1, \cdots, \alpha_n).$$

Hence $p_1 = p_2$ by I. It follows that $f(p)$ is one-to-one on S to $f(S) \subset S^*$. Consider $p \in S$ and $U^*_{\alpha_1 \cdots \alpha_n}(f(p))$. If for some $\{\beta_1, \cdots, \beta_m\} \subset A$

$$\Pi^Q(\beta_1, \cdots, \beta_m) \subset \Pi^P(\alpha_1, \cdots, \alpha_n) = \prod_{i=1}^{n} U_{\alpha_i}(p),$$

then $Q \in U^*_{\alpha_1 \cdots \alpha_n}(f(p))$. Applying induction to II, we obtain an $\alpha = \alpha(\alpha_1, \cdots, \alpha_n ; p)$ such that

$$U_\alpha(p) \subset \prod_{i=1}^{n} U_{\alpha_i}(p).$$

For α and p let $\lambda(\alpha)$, $\delta(p, \alpha)$ be the indices of III. If $q \in U_{\lambda(\alpha)}(p)$, then $U_{\delta(p,\alpha)}(q)U_{\lambda(\alpha)}(p) \neq 0$ so that from III $U_{\delta(p,\alpha)}(q) \subset U_\alpha(p)$ and

$$\Pi^q(\delta(p, \alpha)) = U_{\delta(p,\alpha)}(q) \subset U_\alpha(p) \subset \Pi^p(\alpha_1, \cdots, \alpha_n).$$

Thus

$$f(q) = C(\Pi^q) \subset U^*_{\alpha_1 \cdots \alpha_n}(f(p)),$$

and

$$f[U_{\lambda(\alpha)}(p)] \subset U^*_{\alpha_1 \cdots \alpha_n}(f(p)).$$

This establishes the continuity of f on S to S^*. Conversely, consider $f^{-1}(P) = p \in S$ and $U_\alpha(p)$. If $Q \in U_\alpha^*(P)f(S)$, then $f^{-1}(Q) = q \in S$ and, for some $\beta_i = \beta_i(\alpha, P) \in A$ $(i = 1, \cdots, n)$, $U^*_{\beta_1 \cdots \beta_n}(Q) \subset U^*_\alpha(P)$. That is,

$$q \in \Pi^q(\beta_1, \cdots, \beta_n) \subset \Pi^p(\alpha) = U_\alpha(p).$$

Thus

$$f^{-1}[U_\alpha^*(f(p))f(S)] \subset U_\alpha(p), \qquad p \in S, \alpha \in A$$

and f^{-1} is uniformly continuous on $f(S) \subset S^*$ to S.

Finally consider a Cauchy sequence $p_k \in S$. For each $\alpha \in A$ there are $p_\alpha \in S$ and $k_\alpha > 0$ such that

$$p_k \in U_\alpha(p_\alpha), \qquad k \geqq k_\alpha .$$

For any $\{\alpha_1, \cdots, \alpha_n\} \subset A$ let $k(\alpha_1, \cdots, \alpha_n) = \max_{1 \leq i \leq n} k_{\alpha_i}$ and $k \geqq k(\alpha_1, \cdots, \alpha_n)$. Then

$$p_k \in \prod_{i=1}^n U_{\alpha_i}(p'_{\alpha_i}).$$

Thus $U_\alpha(p'_\alpha)$ determines a Π and a $Q = C(\Pi) \in S^*$. Let $\Pi^Q = U_\alpha(q_\alpha)$. Since $\Pi \sim \Pi^Q$, there is for every $\{\alpha_1, \cdots, \alpha_n\} \subset A$ a set $\{\beta_1, \cdots, \beta_m\} \subset A$, according to Theorem 4, such that

$$\Pi(\beta_1, \cdots, \beta_m) \subset \Pi^Q(\alpha_1, \cdots, \alpha_n)$$

and

$$p_k \in \Pi^Q(\alpha_1, \cdots, \alpha_n), \qquad k \geqq k(\beta_1, \cdots, \beta_m) = K(\alpha_1, \cdots, \alpha_n),$$

since the $\beta_i = \beta_i(\alpha_1, \cdots, \alpha_n)$. Now for each $\alpha \in A$, $U_{\lambda(\alpha)}(q_{\lambda(\alpha)}) \subset U_\alpha(q_{\lambda(\alpha)})$, so that for $k \geqq K(\lambda(\alpha_1), \cdots, \lambda(\alpha_n))$

$$p_k \in \Pi^Q(\lambda(\alpha_1), \cdots, \lambda(\alpha_n)) = \prod_{i=1}^n U_{\lambda(\alpha_i)}(q_{\lambda(\alpha_i)}) \subset \prod_{i=1}^n U_{\alpha_i}(q_{\lambda(\alpha_i)}) \tag{11}$$

for all $\{\alpha_1, \cdots, \alpha_n\} \subset A$. Thus the family $U_\alpha(q_{\lambda(\alpha)})$ defines a $\tilde{\Pi}$. Since

$$0 \neq \Pi^Q(\lambda(\alpha_1), \cdots, \lambda(\alpha_n), \alpha_1, \cdots, \alpha_n) = \prod_{i=1}^n U_{\lambda(\alpha_i)}(q_{\lambda(\alpha_i)}) \prod_{i=1}^n U_{\alpha_i}(q_{\alpha_i})$$

$$\subset \prod_{i=1}^n U_{\alpha_i}(q_{\lambda(\alpha_i)}) \prod_{i=1}^n U_{\alpha_i}(q_{\alpha_i}) = \tilde{\Pi}(\alpha_1, \cdots, \alpha_n)\Pi^Q(\alpha_1, \cdots, \alpha_n),$$

$\tilde{\Pi} \sim \Pi^Q$ by Theorem 2. Again by Theorem 4, there is for every $\{\alpha_1, \cdots, \alpha_n\} \subset A$ a set $\{\gamma_1, \cdots, \gamma_k\} \subset A$ such that

$$\tilde{\Pi}(\gamma_1, \cdots, \gamma_k) \subset \Pi^Q(\alpha_1, \cdots, \alpha_n). \tag{12}$$

For these γ_i consider $\lambda(\gamma_i)$ the first index of III, $q_{\lambda(\gamma_i)}$ and $\delta_i = \delta(q_{\lambda(\gamma_i)}, \gamma_i)$ the second index of III. From (11)

$$p_k \in \Pi^Q(\lambda(\gamma_1), \cdots, \lambda(\gamma_k)) = \prod_{i=1}^k U_{\lambda(\gamma_i)}(q_{\lambda(\gamma_i)}), \qquad k \geqq K(\lambda(\gamma_1), \cdots, \lambda(\gamma_k)).$$

For such k

$$U_{\delta_i}(p_k)U_{\lambda(\gamma_i)}(q_{\lambda(\gamma_i)}) \neq 0.$$

Thus from III and $\delta_i = \delta(q_{\lambda(\gamma_i)}, \gamma_i)$

$$U_{\delta_i}(p_k) \subset U_{\gamma_i}(q_{\lambda(\gamma_i)}) \qquad (i = 1, \cdots, k). \tag{13}$$

Hence from the definition of $\bar{\Pi}$, (12), and (13)

$$\Pi^{p_k}(\delta_1, \cdots, \delta_k) = \prod_{i=1}^{k} U_{\delta_i}(p_k) \subset \prod_{i=1}^{k} U_{\gamma_i}(q_{\lambda(\gamma_i)})$$

$$= \bar{\Pi}(\gamma_1, \cdots, \gamma_k) \subset \Pi^{Q}(\alpha_1, \cdots, \alpha_n),$$

$$k \geqq K(\lambda(\gamma_1), \cdots, \lambda(\gamma_k)) = K^*(\alpha_1, \cdots, \alpha_n).$$

Thus

$$f(p_k) \in U^*_{\alpha_1 \cdots \alpha_n}(Q), \qquad k \geqq K^*(\alpha_1, \cdots, \alpha_n),$$

$$\lim_k f(p_k) = Q.$$

THEOREM 11. *$f(S)$ is dense in S^*.*

Proof. Let $U^*_{\alpha_1 \cdots \alpha_n}(P)$ be any neighborhood in the space S^* and let $\Pi^P = \{U_\alpha(p_\alpha)\}$. It follows from Theorem 4 that $\{U_\alpha(p_{\lambda(\alpha)})\}$ defines a $\Pi' \sim \Pi^P$ and that for the given $\{\alpha_1, \cdots, \alpha_n\}$ there is a set $\{\beta_1, \cdots, \beta_m\} \subset A$ such that

$$\Pi'(\beta_1, \cdots, \beta_m) \subset \Pi^P(\alpha_1, \cdots, \alpha_n).$$

There is a q such that

$$q \in \Pi^P(\lambda(\beta_1), \cdots, \lambda(\beta_m)) = \prod_{i=1}^{m} U_{\lambda(\beta_i)}(p_{\lambda(\beta_i)}),$$

where $\lambda(\beta_i)$ is the first index of III. If $\delta_i = \delta(p_{\lambda(\beta_i)}, \beta_i)$ is the second index of III, then

$$U_{\delta_i}(q) \subset U_{\beta_i}(p_{\lambda(\beta_i)}) \qquad (i = 1, \cdots, m).$$

Hence

$$\Pi^q(\delta_1, \cdots, \delta_m) = \prod_{i=1}^{m} U_{\delta_i}(q) \subset \prod_{i=1}^{m} U_{\beta_i}(p_{\lambda(\beta_i)}) = \Pi'(\beta_1, \cdots, \beta_m) \subset \Pi^P(\alpha_1, \cdots, \alpha_n).$$

Thus

$$f(q) \in U^*_{\alpha_1 \cdots \alpha_n}(P).$$

Remark on the first denumerability axiom. If the set A of indices is the set $(1, 2, \cdots, n, \cdots)$, S and S^* satisfy Hausdorff's first denumerability axiom. In this case we have the following result:

THEOREM 12. *S^* is a space of the second category.*

Proof. It is the case in general that every regular, complete Hausdorff space satisfying the first denumerability axiom is of the second category.

The question arises as to whether a Hausdorff space, satisfying the first denumerability axiom and

III_ω. To each $p \in S$ and $n > 0$ there are positive integers $m(n)$ and $k(p, n)$ such that if $U_{k(p,n)}(q)U_{m(n)}(p) \neq 0$, then $U_{k(p,n)}(q) \subset U_n(p)$,

is metrizable.

Bibliography

P. Alexandroff and P. Urysohn, Comptes Rendus, vol. 177(1923), p. 1274.

E. W. Chittenden and A. D. Pitcher, *On the theory of developments of an abstract space*, Trans. Amer. Math. Soc., vol. 20(1919), pp. 213–233.

L. W. Cohen, *Uniformity properties in topological space*, this Journal, vol. 3(1937), pp. 610–615.

V. W. Niemytski, *On the "Third Axiom of Metric Space"*, Trans. Amer. Math. Soc., vol. 29 (1927), pp. 507–513.

A. Weil, *Sur les Espaces à Structure Uniforme*, Paris, 1938.

Brown University.

THE JACOBI CONDITION AND THE INDEX THEOREM IN THE CALCULUS OF VARIATIONS

By E. J. McShane

In 1916 Bliss[1] published a paper on the Jacobi condition in which he presented the method now familiar to every student of the calculus of variations. For a parametric problem $\int f(y_1, \cdots, y_n, y_1', \cdots, y_n')dt = \text{min.}$ conjugate points on an extremal $y = y(t)$ were defined as zeros of a non-identically vanishing "normal" solution $\eta(t)$ of the Jacobi equations; that is, a solution such that $y_i'(t)\eta_i(t) \equiv 0$.

Hestenes[2] recently commented that, since there are only $2n - 2$ linearly independent normal solutions of the Jacobi equations, Bliss' treatment is quite different from that usually given (i.e., since 1916) for non-parametric problems. He then gave a discussion of conjugate points by adjoining the equation $y_i'\eta_i' = \text{constant}$ (or $y_i'\eta_i'' + y_i''\eta_i' = 0$) to the Jacobi equations. There are $2n$ linearly independent solutions of the resulting system of equations.

More recently, Birkhoff and Hestenes[3] used the device of adjoining any one of several second-order equations to the Jacobi equations in studying natural isoperimetric conditions, and Morse[4] used one of these equations

$$f_{y^i}\eta^i + f_{y^{i'}}\eta^{i'} = \text{const.}$$

(assuming $f > 0$) in establishing the index theorem.

As compared with earlier treatments of the Jacobi condition, the advantages of Bliss' method are too well known to require mention. But even in comparison with the more recent treatments just mentioned it has the advantage that the property of being a "normal" solution is evidently invariant under change of parameter along the extremal, so that all representations of the curve are equally usable. This is not true of the "special" solutions of Hestenes,[5] and the treatment of Morse rests on a special choice of parameter.

Received January 16, 1939.

[1] G. A. Bliss, *Jacobi's condition for problems of the calculus of variations in parametric form*, Trans. Am. Math. Soc., vol. 17(1916), pp. 195–206.

[2] M. R. Hestenes, *A note on the Jacobi condition* ⋯, Bull. Am. Math. Soc., vol. 40(1934), pp. 297–302.

[3] G. D. Birkhoff and M. R. Hestenes, *Natural isoperimetric conditions in the calculus of variations*, this Journal, vol. 1(1935), p. 264.

[4] M. Morse, *The index theorem in the calculus of variations*, this Journal, vol. 4(1938), pp. 231–246.

[5] However, it should be remarked that Hestenes suggests several other equally useful forms of adjoined equation, and one of these ($y_i'\eta_i'/y_j'y_j' = \text{constant}$) gives solutions with the invariance property under discussion.

Furthermore, Bliss' statement of the parametric problem can be made, by a simple device, to include the non-parametric case. While the non-parametric problem offers less difficulty than the parametric in so far as the Jacobi condition is concerned, it seems a definite gain to be able thus to obtain the results in one case practically free of charge.

In this note I present a discussion of conjugate points which, being only a minor modification of Bliss' method, retains its advantages, and moreover, (1) is invariant under change of parameter and of coördinates, and (2) can be used as well as Morse's for proving the index theorem. Another advantage is that the *tests* for conjugate points in non-parametric problems follow at once from the tests developed for parametric problems.

2. Throughout this paper y and r will denote either n-tuples $(y^1, \cdots, y^n)$, $(r^1, \cdots, r^n)$ or $(n+1)$-tuples $(y^0, \cdots, y^n)$, $(r^0, \cdots, r^n)$. The danger of confusion will be diminished by using β and (except in §§13, 14) α as variables on the range $0, 1, \cdots, n$ and i and j as variables on the range $1, \cdots, n$. Likewise η and ρ will denote n-tuples or $(n+1)$-tuples according to need, the superscript indicating the number of components. The principal object of study is the integral

$$J(C) = \int_{t_1}^{t_2} F(y(t), \dot{y}(t))\, dt, \tag{2.1}$$

an expression whose elements we must discuss more fully. In order to have adequate generality, we assume that y is a point of a manifold M. There may be no single coördinate system for all of M, but it will be assumed that every point of M has neighborhoods homeomorphic with a Euclidean n-sphere, so that local coördinates can be introduced. We assume that M can be covered by local systems in such a manner that if y, $\bar{y}$ are the coördinates in two local systems of a common point, then the y^i are functions of class C^4 of the $\bar{y}^i$, and conversely.

Let us then suppose that R is a collection of elements (y, r), where y ranges over a point set in M, and for each y the element r ranges over a set of contravariant vectors such that if (y, r) is in R, so is (y, kr) for all $k > 0$. We make the following assumptions concerning the integrand. It is defined and continuous on R. At each point (y, r) of R with $|r| > 0$, the function F has continuous partial derivatives of all orders up to and including the fourth. It is positively homogeneous of degree 1 in r. Under a change of coördinates from y to $\bar{y}$ which replaces the contravariant vector r by $\bar{r}$, the function $F(y, r)$ is invariant; that is, it is replaced by $\bar{F}(\bar{y}, \bar{r}) = F(y, r)$.

Well-known consequences of the positive homogeneity

$$F(y, kr) = kF(y, r) \qquad (k > 0) \tag{2.2}$$

of F are the identities

$$r^i F_{r^i}(y, r) = F(y, r), \tag{2.3}$$

$$r^i F_{r^i r^j}(y, r) = 0. \tag{2.4}$$

The curves C considered will be those of class D^1; that is, those curves having representations

$$C: y = y(t) \qquad (t_1 \leqq t \leqq t_2) \tag{2.5}$$

in which each function $y^i(t)$ is of class D^1 (i.e., continuous and with a derivative which has at most a finite number of simple jump discontinuities) and $y'^1 y'^2 \cdots y'^n$ is bounded from zero. If $t(\tau)$ is a function of class D^1 with $t'(\tau) > 0$ $(\tau_1 \leqq \tau \leqq \tau_2)$ which satisfies $t(\tau_s) = t_s$ $(s = 1, 2)$, then

$$y = \eta(\tau) \equiv y(t(\tau)) \qquad (\tau_1 \leqq \tau \leqq \tau_2) \tag{2.6}$$

is another representation of C.

The curve C is *admissible* if it has a representation (2.5) such that $(y(t), y'(t))$ is in R for $t_1 \leqq t \leqq t_2$. It follows readily that if C is admissible and (2.6) is also a representation of C, then $(\eta(\tau), \eta'(\tau))$ is in R for $\tau_1 \leqq \tau \leqq \tau_2$. Thus the two integrals

$$\int_{t_1}^{t_2} F(y, y')\, dt, \qquad \int_{\tau_1}^{\tau_2} F(\eta, \eta')\, d\tau$$

are both defined. By the usual proof, the positive homogeneity of F implies the equality of the two integrals; we denote their common value by $J(C)$.

As usual, we say that an admissible curve C gives a weak relative minimum to $J(C)$ in a class $\mathcal{K}$ of admissible curves if there is a representation (2.5) of C and a positive number ϵ such that the inequality $J(\bar{C}) \geqq J(C)$ holds for every curve $\bar{C}$ of $\mathcal{K}$ which has a representation

$$\bar{C}: y = \bar{y}(t) \qquad (t_1 \leqq t \leqq t_2) \tag{2.7}$$

such that

$$|\bar{y}(t) - y(t)| < \epsilon \qquad (t_1 \leqq t \leqq t_2) \tag{2.8}$$

and

$$|\bar{y}(t) - y'(t)| < \epsilon \tag{2.9}$$

except at corners. If (2.8) holds, it continues to hold under any simultaneous change of parameter on C and $\bar{C}$. The same is not true of (2.9). But it is easily seen that if C gives a weak relative minimum to $J(C)$, then to *each* representation of C there corresponds a positive ϵ such that $J(\bar{C}) \geqq J(C)$ subject to (2.8) and (2.9). Henceforward we shall be interested in the properties of a curve $C: y = y(t)$ $(t_1 \leqq t \leqq t_2)$ which gives a weak relative minimum to $J(C)$ and has $(y(t), y'(t))$ interior to R for $t_1 \leqq t \leqq t_2$.

By the usual methods we can show that for such a curve C the Weierstrass-Erdmann corner condition

$$F_{y'^i}(y(t), \dot{y}(t - 0)) = F_{y'^i}(y(t), \dot{y}(t + 0)) \qquad (i = 1, \cdots, n) \tag{2.10}$$

holds for $t_1 < t < t_2$, while the Euler equations

$$L_i(y) \equiv \frac{d}{dt} F_{\dot{y}^i}(y, \dot{y}) - F_{y^i}(y, \dot{y}) = 0 \qquad (i = 1, \cdots, n) \tag{2.11}$$

hold except perhaps at corners of C. These equations are not independent; they satisfy the identity

$$\dot{y}^i L_i(y) \equiv 0, \tag{2.12}$$

as follows from (2.3) and (2.4).

A curve of class C^2 which satisfies the Euler equations is called an extremal. A curve C, represented by (2.5), is non-singular if the determinant

$$| F_{\dot{y}^i \dot{y}^j}(y(t), y'(t)) | \tag{2.13}$$

(which must be zero because of (2.4)) has rank $n - 1$. By a theorem of Hilbert, a non-singular curve of class C^1 which satisfies the Euler equations can be given a representation (2.5) in which the functions $y^i(t)$ are of class C^4; in fact, this holds if arc length is chosen as parameter. Henceforth we adopt the standing hypothesis and notation that *the curve*

$$g: \; y^i = \gamma^i(t) \qquad (t_1 \leqq t \leqq t_2) \tag{2.14}$$

is a non-singular extremal, the functions $\gamma^i(t)$ *being of class* C^4 *and* $(\gamma, \dot{\gamma})$ *being interior to* R *for* $t_1 \leqq t \leqq t_2$. For compactness, we use the superscript 0 on F or any of its partial derivatives to denote the arguments on $(\gamma(t), \dot{\gamma}(t))$. It will be observed that the notation is chosen as though it were possible to choose a single coördinate system for a neighborhood of the whole of g. This is not essential for (2.10) and (2.11), which are local properties; but it is very useful in studying the Jacobi condition. That there is no loss of generality in this assumption has been shown by Morse.[6]

Let $\eta(t)$ be a function of class D^1 which vanishes at t_1 and t_2. The second variation of $J(C)$ along γ due to the variation η is

$$I(\eta) \equiv \int_{t_1}^{t_2} 2\Omega(t, \eta, \dot{\eta})\, dt, \tag{2.15}$$

where[7]

$$2\Omega(t, \eta, \rho) = F^0_{x^i x^j} \eta^i \eta^j + 2F^0_{x^i \dot{y}^j} \eta^i \rho^j + F^0_{\dot{y}^i \dot{y}^j} \rho^i \rho^j. \tag{2.16}$$

If the curve g gives a weak relative minimum to $J(C)$ in the class of curves joining the end points of g, this second variation must be non-negative. The Jacobi equations for $J(C)$ are the Euler equations for $I(\eta)$, and have the form

$$J_i(\eta) \equiv \frac{d}{dt} \Omega_{\rho^i} - \Omega_{\eta^i} = 0. \tag{2.17}$$

[6] M. Morse, *The Calculus of Variations in the Large*, Am. Math. Soc. Colloquium Publications, vol. XVII, p. 108.

[7] We depart from traditional notation in writing ρ in place of $\dot{\eta}$. This is in analogy with the notation $F(y, r)$ and avoids printing complicated subscripts in (2.17) and later equations involving Ω_{ρ^i}.

The problem in non-parametric form is concerned with the integral

$$J[y] = \int_{x_1}^{x_2} f(x, y(x), \dot{y}(x))\, dx. \tag{2.18}$$

Here $f(x, y, r)$ is defined and of class C^4 for all (x, y, r) in a set R^* in $(2n + 1)$-space.

We avoid a separate discussion of this problem by the following device. Let R be the set of all points $(y^0, y^1, \cdots, y^n, r^0, r^1, \cdots, r^n)$ such that $r^0 > 0$ and

$$(y^0, y^1, \cdots, y^n, r^1/r^0, \cdots, r^n/r^0)$$

is in R^*. On R we define

$$F(y, r) = r^0 f(y^0, y^1, \cdots, y^n, r^1/r^0, \cdots, r^n/r^0). \tag{2.19}$$

This has the properties required of F in the preceding paragraph, y and r being now $(n + 1)$-tuples. A curve $y^i = y^i(x)$ $(x_1 \leqq x \leqq x_2)$ can be regarded as a curve C: $y^\alpha = y^\alpha(t)$ $(x_1 \leqq t \leqq x_2)$ in which $y^0(t) \equiv t$. From the identity

$$F(y^0, \cdots, y^n, 1, r^1, \cdots, r^n) = f(y^0, \cdots, y^n, r^1, \cdots, r^n) \tag{2.20}$$

derived from (2.19) we see that for this representation of C (and therefore for all representations of C) we have $J(C) = J[y]$. Thus for every curve C in R which has a D′ representation of the form $y = y(x)$ the functionals $J(C)$ and $J[y]$ differ only in that the special representation $y = y(x)$ is needed to compute $J[y]$, whereas the same number $J[y]$ is arrived at by computing $J(C)$ for an arbitrary representation of C. This makes it evident that the problem of finding necessary and sufficient conditions that a curve C: $y = y(x)$ minimize $J[y]$ is equivalent to the problem of finding such conditions when $J[y]$ is replaced by $J(C)$. The only difficulty, and that a minor one, is that the conditions thus arrived at are expressed in terms of the auxiliary function F rather than in terms of f itself, and it is desirable to reduce them to such a form that only f and its derivatives enter. This is quite easy to do. From the identity (2.20) we obtain by differentiation a set of similar identities for all the existing partial derivatives which do not involve differentiation with respect to r^0. Since F is homogeneous, equation (2.3) holds, and for $r^0 = 1$ this yields

$$F_{r^0}(y, r) = F - r^i F_{r^i} = f - r^i f_{r^i}. \tag{2.21}$$

The Euler equations for $J(C)$ then take the form

$$\frac{d}{dx} f_{r^i} - f_{y^i} = 0, \qquad \frac{d}{dx}(f - y^i f_{r^i}) - f_x = 0. \tag{2.22}$$

The last of these is a consequence of the others, by (2.12). If we represent g by equations $y^\alpha = \gamma^\alpha(t)$ with $\gamma^0(t) \equiv t$, the matrix (2.13) takes the form

$$|F^0_{r^\alpha r^\beta}| = \begin{vmatrix} F^0_{r^0 r^0} & F^0_{r^0 r^j} \\ F^0_{r^i r^0} & f^0_{r^i r^j} \end{vmatrix}. \tag{2.23}$$

If $C: y = y(x)$ is such that

$$|f^0_{r^i r^j}| \neq 0,$$

it is non-singular by the definition above. The converse is also true, since the first row [column] of (2.23) is a linear combination of the other rows [columns] by (2.4).

The Jacobi equations for $J[y]$ are by definition

$$J_i(\eta) \equiv \frac{d}{dx}\omega_{\rho^i} - \omega_{\eta^i} = 0, \tag{2.24}$$

where

$$2\omega(x, \eta, \rho) = f^0_{y^i y^j}\eta^i\eta^j + 2f^0_{y^i r^j}\eta^i\rho^j + f^0_{r^i r^j}\rho^i\rho^j. \tag{2.25}$$

Equations (2.24) differ from the Jacobi equations for $J(C)$ only in that every term involving a partial derivative of F as to r^0 or y^0 is omitted. If we observe that every such term has a factor η^0 or $\dot{\eta}^0$, we have at once

LEMMA 2.1. *The functions* $(\eta^1(t), \cdots, \eta^n(t))$ *form a solution of the Jacobi equations* (2.24) *for* $J[y]$ *if and only if the functions* $(0, \eta^1(t), \cdots, \eta^n(t))$ *form a solution of the Jacobi equations* (2.17) *for* $J(C)$.

3. At this point I wish to permit myself the luxury of a digression on the organization of the subject matter of the elementary calculus of variations, from a pedagogical point of view. There is a great similarity between the parametric and non-parametric problems, so much in fact that a sequence of lectures covering both problems frequently contains a considerable amount of repetition required by the need of changing a few details. If, however, the parametric problem is treated in full detail, with the assumption as above that $F(y, r)$ is defined for all (y, r) in a set R (not necessarily containing all r), essentially no added effort is required in discussing the parametric problem, and the results for non-parametric problems fall out at once. This is especially true if strong and weak relative minima are defined as above in terms of the *curves* defined by $y = y(x)$ instead of in terms of the *functions* $y(x)$. (Incidentally, this avoids some verbosity, especially in defining weak relative minima for problems involving variable end points or involving minimizing curves which are not of class C^1.) The Euler equations being established for the parametric form, we need only introduce x as parameter to obtain (2.22). The Weierstrass condition and Legendre conditions are likewise obtained. For example, let it have been proved that

$$u^\alpha F^0_{r^\alpha r^\beta} u^\beta \geqq 0 \tag{3.1}$$

for all u. If in particular we set $u^0 = 0$, we have

$$u^i f^0_{r^i r^j} u^j \geqq 0,$$

as a consequence of (2.23). Similarly, if the Jacobi condition is established for $J(C)$, it holds for the non-parametric problem. All that remains is to express the condition in terms of f, and this we shall do in §11.

A similar situation holds in sufficiency theorems. For parametric integrals with a general range R for (y, r) we find that the Euler condition, the strengthened Legendre condition, the strengthened Jacobi condition and the Weierstrass condition in the doubly strengthened form (II_b', in the notation of Bliss[8]) are sufficient for a strong relative minimum. It is easy to verify that if $y = \gamma(x)$ satisfies the analogous conditions in terms of $f(x, y, r)$, then it satisfies these conditions for the associated parametric integral. For the usual type of parametric problem it can be shown as usual that II' and the strengthened Legendre condition imply II_b'.

4. We return to our problem. For parametric problems the Jacobi equations are not independent; they satisfy the equation

$$\dot{\gamma}^i J_i(\eta) = 0 \tag{4.1}$$

identically[9] in t, η, $\dot{\eta}$, and $\ddot{\eta}$. This leads to a superfluity of solutions of the Jacobi conditions. In fact, assuming as usual that

$$g\colon\ y = \gamma(t) \qquad (t_1 \leqq t \leqq t_2) \tag{4.2}$$

is a non-singular extremal, we can prove

LEMMA 4.1. *If $\rho(t)$ is any function of class C^2 on the interval $[t_1, t_2]$, the functions $\rho(t)\dot{\gamma}^i(t)$ satisfy the Jacobi equations.*

The extremal g can be extended so as to be defined on an interval $t_1 - \epsilon \leqq t \leqq t_2 + \epsilon$, $\epsilon > 0$. For all small e the curve

$$y^i = y^i(t, e) \equiv \gamma^i(t + e\rho(t)) \qquad (t_1 \leqq t \leqq t_2) \tag{4.3}$$

is an arc of this extended extremal g, hence is an extremal. By Jacobi's theorem, since (4.3) is an extremal for all small e, the partial derivatives

$$y_e^i(t, 0) = \dot{\gamma}^i(t)\rho(t) \tag{4.4}$$

satisfy the Jacobi equations.

We therefore cannot profitably say that two points $\gamma(t_3)$, $\gamma(t_4)$ are conjugate if they are zeros of a non-identically vanishing solution of the Jacobi equations, for by this definition any two points of g would be conjugate. It is necessary to select in some manner among the solutions of the Jacobi equations. This selection can be effected in an assortment of ways. We shall select a class which we shall call p-normal solutions.

5. Let us suppose that, given the representation (4.2) of g, to each t in $[t_1, t_2]$ there corresponds a vector $p(t) \equiv (p_1(t), \cdots, p_n(t))$ with the following properties:

[8] See, e.g., Bliss' Carus Monograph, *The Calculus of Variations*, pp. 135 and 158.
[9] Bliss, loc. cit. (footnote 1), p. 201.

(5.1) *Each $p_i(t)$ is of class* C^2.

(5.2) *The inequality*

$$p_i(t)\dot{\gamma}^i(t) \neq 0$$

holds for $t_1 \leqq t \leqq t_2$.

(5.3)[10] *For each fixed coördinate system, if the change of parameter* $t = t(\tau)$ *(of class* C^2*) is made on* g, *the vector* $p_i(t)$ *is replaced by* $q_i(\tau)$, *where*

$$q_i(\tau) = k(\tau)p_i(t(\tau)),$$

the function $k(\tau)$ *being non-vanishing and of class* C^2.

(5.4)[11] *For each fixed representation of* g *and each fixed value of* t, *the vector* $p(t)$ *is a covariant vector under change of coördinates.*

For example, if the space has a Riemannian metric defined by $g_{ij}(y)$, we can choose

$$p_i(t) = g_{ij}(y_0(t))\dot{\gamma}^j(t).$$

If $F^0 > 0$, then

$$p_i(t) = F_{\tau^i}(\gamma(t), \dot{\gamma}(t))$$

satisfies the conditions. The functions

$$p_i(t) = \dot{\gamma}^i(t)$$

satisfy (5.1), (5.2) and (5.3), but not (5.4).

A function $\eta(t)$ will be called *p-normal* if it satisfies the equation

$$p_i(t)\dot{\eta}^i(t) \equiv 0 \qquad (t_1 \leqq t \leqq t_2). \tag{5.5}$$

To save repetition, we shall usually use the phrase "p-normal solution" as an abbreviation for "p-normal solution of the Jacobi equations (2.17)".

Since the Jacobi equations and the equations (5.5) are linear, it is evident that every linear combination of p-normal solutions is again a p-normal solution.

6. In order to make use of p-normality we establish some lemmas.

LEMMA 6.1. *For each* t *in the interval* $[t_1, t_2]$, *the only solution* $c^1, \cdots, c^n$ *of the equations*[12]

$$F^0_{\tau^i\tau^j}c^j = 0, \qquad p_j(t)c^j = 0 \tag{6.1}$$

is $c^1 = \cdots = c^n = 0$.

Let the c^i satisfy (6.1). The first n of these equations hold with $\dot{\gamma}^j(t)$ in place of c^j, by (2.4). The determinant of coefficients of these n equations is (2.13) with $\gamma(t)$ in place of $y(t)$, and this has rank $n - 1$ since the extremal g

[10] This condition is needed only to establish invariance under change of parameter.

[11] This condition is needed only to establish invariance under change of coördinates.

[12] Recall that the superscript 0 indicates that the arguments are $(\gamma(t), \dot{\gamma}(t))$.

is non-singular. Hence any two solutions of these equations are linearly dependent, and there is a number k such that

$$c^i = k\dot{\gamma}^i(t). \tag{6.2}$$

Substituting this in the last of equations (6.1) and recalling (5.2), we have $k = 0$. This with (6.2) establishes the lemma.

LEMMA 6.2. *The determinant*

$$\begin{vmatrix} F^0_{r'i'rj} & p_i \\ p_j & 0 \end{vmatrix} \tag{6.3}$$

is not zero.

If it were, there would be numbers $c^1, \cdots, c^n, d$ not all zero such that

$$F^0_{r'i'rj}c^j + dp_i = 0 \qquad (i = 1, \cdots, n),$$

$$p_jc^j = 0.$$

Multiplying the first n of these by $\dot{\gamma}^i$ and summing yield (with (2.4))

$$dp_i\dot{\gamma}^i = 0;$$

so by (5.2) we see that $d = 0$. But now the c^j are non-zero solutions of equations (6.1), and this is impossible by Lemma 6.1.

7. It is evident that every p-normal solution η satisfies the set of equations

$$J_i(\eta) + \lambda p_i = 0, \qquad (p_i\eta^i)'' = 0 \tag{7.1}$$

with $\lambda = 0$. On the other hand, if η, λ satisfy (7.1), then λ vanishes identically, as we see by multiplying the first n equations by $\dot{\gamma}^i$ and summing, and using (4.1) and (5.2). From the definition (2.17) of $J_i(\eta)$ we see that the determinant of coefficients of the variables

$$\ddot{\eta}^1, \cdots, \ddot{\eta}^n, \lambda \tag{7.2}$$

in (7.1) is the determinant (6.3), which is not zero. Hence equations (7.1) can be solved for the variables (7.2), and (disregarding the equation for λ) we obtain a system of equations

$$\ddot{\eta}^i = A_{ij}\dot{\eta}^j + B_{ij}\eta^j \tag{7.3}$$

equivalent (when the equation $\lambda \equiv 0$ is added) to (7.1). We can therefore state

LEMMA 7.1. *A set of functions $\eta(t) \equiv (\eta^1(t), \cdots, \eta^n(t))$ satisfies (7.3) if and only if $\eta(t)$ satisfies the Jacobi equations and $p_i(t)\eta^i(t)$ is a linear function of t. In particular, every p-normal solution satisfies (7.3); and if $\eta(t)$ is a solution of (7.3) such that $p_i\eta^i$ and $(p_i\eta^i)'$ both vanish for some value of t, then $p_i\eta^i$ vanishes identically, and $\eta(t)$ is a p-normal solution.*

From the well-known theorem on solutions of linear differential equations we have

LEMMA 7.2. *If η_1 and η_2 are two solutions of equations (7.3) which at some value t_3 of t are equal and have equal derivatives, they are identically equal. In particular, if η is a p-normal solution which vanishes with its derivative at some value t_3 of t, it is identically zero.*

The last remark follows from the fact that η and $\eta_2 \equiv 0$ both satisfy (7.3), and are equal and have equal derivatives at t_3.

8. Our definition of conjugacy is as follows:

If $\gamma(t_3)$ and $\gamma(t_4)$ are two points of the extremal g $(t_1 \leqq t_3 < t_4 \leqq t_2)$ such that there exists a p-normal solution $\eta(t)$ which vanishes at t_3 and t_4 but is not identically zero, then these two points are conjugate to each other.

The Jacobi necessary condition is

THEOREM 8.1. *If g is a non-singular extremal which gives a weak relative minimum to the integral $J(C)$, then there is no number t_3 in the interval $t_1 < t < t_2$ such that $\gamma(t_3)$ is conjugate to $\gamma(t_1)$.*

Suppose the contrary. There is a p-normal solution $\eta(t)$ not identically zero which vanishes at t_1 and at t_3, where $t_1 < t_3 < t_2$. Define

$$
\begin{aligned}
u^i(t) &= \eta^i(t) && (t_1 \leqq t \leqq t_3), \\
u^i(t) &\equiv 0 && (t_3 < t \leqq t_2).
\end{aligned} \tag{8.1}
$$

Then $u(t)$ is of class D^1. It yields the second variation

$$
I(u) = \int_{t_1}^{t_2} 2\Omega(t, u, \dot{u})\, dt = \int_{t_1}^{t_3} 2\Omega(t, \eta, \dot{\eta})\, dt. \tag{8.2}
$$

Since Ω is homogeneous quadratic in η and ρ, it satisfies the relation

$$
2\Omega(t, \eta, \dot{\eta}) = \eta^i \Omega_{\eta^i} + \dot{\eta}^i \Omega_{\rho^i}. \tag{8.3}
$$

Substituting in (8.2) and integrating by parts give

$$
I(u) = \eta^i \Omega_{\rho^i} \Big|_{t_1}^{t_3} + \int_{t_1}^{t_3} \eta^i \left(\Omega_{\eta^i} - \frac{d}{dt} \Omega_{\rho^i} \right) dt = 0, \tag{8.4}
$$

since η vanishes at t_1 and t_3 and satisfies equations (2.17). But $I(\eta)$, being the second variation of our integral, is never negative; so $u(t)$ minimizes $I(\eta)$ in the class of all D^1 functions vanishing at t_1 and t_2. It therefore must satisfy the necessary conditions for a minimum, and in particular the Weierstrass-Erdmann corner condition:

$$
\Omega_{\rho^i}(t_3, u(t_3), \dot{u}(t_3 - 0)) = \Omega_{\rho^i}(t_3, u(t_3), \dot{u}(t_3 + 0)). \tag{8.5}
$$

But by (8.1) we have

$$
u^i(t_3) = \dot{u}^i(t_3 + 0) = 0, \qquad \dot{u}^i(t_3 - 0) = \dot{\eta}(t_3) \tag{8.6}
$$

If we recall the definition (2.16) of Ω, equation (8.5) becomes

$$F^0_{r,i,r,j}\eta^j = 0. \tag{8.7}$$

Since η is p-normal, we have on differentiating (5.5)

$$\dot{p}_i(t_3)\eta^i(t_3) + p_i(t_3)\dot{\eta}^i(t_3) = p_i(t_3)\dot{\eta}^i(t_3) = 0. \tag{8.8}$$

Equations (8.7) and (8.8) imply

$$\dot{\eta}^i(t_3) = 0 \qquad (i = 1, \cdots, n) \tag{8.9}$$

by Lemma 6.1. By Lemma 7.2, η vanishes identically. This contradiction establishes the theorem.

9. To develop criteria for conjugate points we need some properties of p-normal solutions. We observe to begin with that there is no loss of generality in supposing that

$$p_i(t)\dot{\gamma}^i(t) \equiv 1. \tag{9.1}$$

For we can replace $p_i(t)$ by

$$p_i(t)/p_i(t)\dot{\gamma}^i(t),$$

by (5.2), without injury to any of the conditions (5.1) to (5.4).

LEMMA 9.1. *To every solution η of the Jacobi equations there corresponds a unique p-normal solution of the form*

$$u^i(t) = \eta^i(t) - \rho(t)\dot{\gamma}^i(t). \tag{9.2}$$

The functions $u(t)$ vanish for all t such that $\eta(t) = 0$.

If we take

$$\rho(t) = p_i(t)\eta^i(t), \tag{9.3}$$

we see immediately that $u(t)$ is p-normal. By Lemma 4.1, it is a solution of the Jacobi equations. Conversely, if $u(t)$ is p-normal, then on multiplying by p_i and using (9.1), we obtain (9.3). So the solution is unique.

LEMMA 9.2. *The functions* H_1, H_2 *defined by*

$$H_1^i(t) = (t - t_1)\dot{\gamma}^i(t), \qquad H_2^i(t) = \dot{\gamma}^i(t) \tag{9.4}$$

satisfy equations (7.3).

By (9.1) we have

$$p_i(t)H_1^i(t) = t - t_1, \qquad p_i(t)H_2^i(t) = 1. \tag{9.5}$$

By Lemma 4.1 the functions $H_1(t)$ and $H_2(t)$ satisfy the Jacobi equations. So by Lemma 7.1 they satisfy equations (7.3).

LEMMA 9.3. *If $\eta_1, \cdots, \eta_{2n-2}$ are linearly independent p-normal solutions, every p-normal solution is a linear combination of them.*

Consider the determinant

$$\begin{vmatrix} \eta_k^i(t) & H_s^i(t) \\ \eta_k^i(t) & \dot{H}_s^i(t) \end{vmatrix} \qquad (k = 1, \cdots, 2n-2;\ i = 1, \cdots, n;\ s = 1, 2). \tag{9.6}$$

Suppose that this vanishes at t_3. Then there are constants c^k, C^s not all zero such that the function

$$c^k \eta_k^i(t) + C^s H_s^i(t) \tag{9.7}$$

vanishes with its derivative at t_3. It then is identically zero by Lemma 7.2. Multiplying by $p_i(t)$ and summing give, since the η_k are p-normal and (9.5) holds,

$$C^1(t - t_1) + C^2 \equiv 0.$$

Hence $C^1 = C^2 = 0$, and the function $c^k \eta_k^i(t)$ vanishes identically. This is impossible, since the η_k are linearly independent.

Since the determinant (9.6) is non-vanishing, every solution of equations (7.3) is a linear combination of its columns. Let $\eta(t)$ be any p-normal solution. It satisfies (7.3), so is a linear combination of the η_k and H_s:

$$\eta^i(t) = c^k \eta_k^i(t) + C^s H_s^i(t). \tag{9.8}$$

Multiplying by p_i and summing give (with (9.5))

$$0 = C^1(t - t_1) + C^2,$$

so $C^1 = C^2 = 0$, and (9.8) expresses η as a linear combination of the η_k.

LEMMA 9.4. *If $\eta_1, \cdots, \eta_{n-1}$ are linearly independent p-normal solutions vanishing at t_1, every p-normal solution vanishing at t_1 is a linear combination of them.*

Suppose that the determinant

$$| \eta_h^i(t_1) \quad \dot{H}_1^i(t_1) | \qquad (h = 1, \cdots, n-1;\ i = 1, \cdots, n) \tag{9.9}$$

is zero. There is then a zero combination of its columns with coefficients c^h, C^1. Therefore the function

$$c^h \eta_h^i(t) + C^1 H_1^i(t) \tag{9.10}$$

vanishes with its derivative at t_1. Since it satisfies (7.3), it vanishes identically, by Lemma 7.2. Multiplying by p_i and summing give (with (9.5))

$$C^1(t - t_1) \equiv 0,$$

so $C^1 = 0$. Therefore the vanishing of (9.10) expresses a linear dependence among the η_h, contrary to hypothesis. We conclude that the determinant (9.9) is not zero.

For every p-normal solution $\eta(t)$ vanishing at t_1 we can find constants c^h, C^1 such that

$$\eta^i(t_1) = c^h \eta_h^i(t_1) + C^1 \dot{\mathrm{H}}_1^i(t_1). \tag{9.11}$$

Thus the two members of the equation

$$\eta^i(t) = c^h \eta_h^i(t) + C^1 \mathrm{H}_1^i(t) \tag{9.12}$$

are equal and have equal derivatives at t_1, and so (9.12) is an identity. Multiplying by p_i and summing give $C^1 = 0$, and $\eta(t)$ is expressed as a linear combination of the $\eta_h(t)$.

10. We now take over §4 of Bliss' paper[13] almost verbatim. It is only necessary to replace the words "normal" by "p-normal", and in the last line on page 203 to replace $y'(t_3)$ by $p(t_3)$. Also, the references to Lemmas 5, 6, 7 are replaced by references to our Lemmas 9.1, 9.3, 9.4, respectively. This gives us

THEOREM 10.1. *If $\eta_1, \cdots, \eta_{2n-2}$ are solutions of the Jacobi equations and the determinant*

$$\Theta(t_1, t) = \begin{vmatrix} \eta_k^i(t) & \dot{\gamma}^i(t) & 0 \\ \eta_k^i(t_1) & 0 & \dot{\gamma}^i(t_1) \end{vmatrix} \qquad (k = 1, \cdots, 2n-1)$$

is not identically zero, then its zeros $t_3 \neq t_1$ determine the points on g conjugate to $\gamma(t_1)$.

THEOREM 10.2. *If $\eta_1, \cdots, \eta_{n-1}$ are solutions of the Jacobi equations such that the determinant*

$$D(t, t_1) = \| \eta_h^i(t) \quad \dot{\gamma}^i(t) \| \qquad (h = 1, \cdots, n-1)$$

has rank 1 at t_1 but is not identically zero, then its zeros $t_3 \neq t_1$ determine the points on g conjugate to $\gamma(t_1)$.

COROLLARY. *If $y^i = y^i(t, \alpha_1, \cdots, \alpha_{n-1})$ $(t_1 \leqq t \leqq t_2)$ is a C^2 family of extremals such that*

$$y^i(t, 0, \cdots, 0) = \gamma^i(t) \qquad (t_1 \leqq t \leqq t_2)$$

and

$$y^i(t_1, \alpha_1, \cdots, \alpha_{n-1}) = \gamma^i(t_1),$$

then either the determinant

$$\Delta(t, t_1) \equiv \frac{\partial(y^1, \cdots, y^n)}{\partial(\alpha_1, \cdots, \alpha_{n-1}, t)}\Bigg|_{\alpha=0}$$

vanishes identically or its zeros $t_3 \neq t_1$ determine the points conjugate to $\gamma(t_1)$.

[13] See footnote 1.

For the first $n-1$ columns of Δ are solutions of the Jacobi equations, by Jacobi's theorem, and the last column is $\dot{\gamma}^i(t)$. The statement then follows from Theorem 10.2.

11. As remarked in §3, we now have a Jacobi condition for the non-parametric integral (2.18) in the form that the extremal $g: y = \gamma(x)$ $(x_1 \leqq x \leqq x_2)$ must not have on it any points conjugate to $\gamma(x_1)$, conjugacy being defined in terms of the associated parametric integrand (2.19). All that remains is to reduce the criteria to terms of the integral (2.18) itself.

Since on our extremal g, written in parametric form

$$y^0 = x = t, \qquad y^i = \gamma(t) \qquad (x_1 \leqq t \leqq x_2) \tag{11.1}$$

we have

$$\dot{y}^0 \equiv 1, \tag{11.2}$$

we can choose $p(t)$ to be the vector $(1, 0, \cdots, 0)$. Conditions (5.1), (5.2) and (5.3) are satisfied. A function $\eta(t)$ is p-normal if and only if

$$\eta^0(t) \equiv 0. \tag{11.3}$$

Therefore by Lemma 2.1 we have

LEMMA 11.1. *The function $(\eta^0(t), \cdots, \eta^n(t))$ is a p-normal solution of the Jacobi equations* (2.17) *for $\int F\,dt$ if and only if $\eta^0 \equiv 0$ and $(\eta^1, \cdots, \eta^n)$ is a solution of the Jacobi equations* (2.24) *for $\int f\,dx$.*

So if we have a set of linearly independent solutions $\eta_1, \cdots, \eta_n$ of the Jacobi equations (2.24), by prefixing a first component 0 to each we obtain n linearly independent p-normal solutions of (2.17).

The definition of conjugacy thus becomes

The points $\gamma(x_3)$, $\gamma(x_4)$ $(x_1 \leqq x_3 < x_4 \leqq x_2)$ are conjugate if there exists a solution η of the (non-parametric) Jacobi equations (2.24) *vanishing at x_3 and x_4 and not identically zero.*

The preceding theorems yield

THEOREM 11.1. *If $\eta_1, \cdots, \eta_{2n}$ are solutions of the Jacobi equations* (2.24) *and the determinant*

$$\theta(x, x_1) = \begin{vmatrix} \eta_k^i(x) \\ \eta_k^i(x_1) \end{vmatrix} \qquad (i = 1, \cdots, n;\ k = 1, \cdots, 2n)$$

is not identically zero, then its zeros $x_3 \neq x_1$ determine the points on E conjugate to $\gamma(x_1)$.

The functions $(0, \eta_k^1, \cdots, \eta_k^n)$ are solutions of the Jacobi equations (2.17). The determinant Θ of Theorem 10.1 takes the form (if we recall (11.2))

$$\begin{vmatrix} 0 & 1 & 0 \\ \eta_k^i(x) & \dot{\gamma}^i(x) & 0 \\ 0 & 0 & 1 \\ \eta_k^i(x_1) & 0 & \dot{\gamma}^i(x_1) \end{vmatrix}.$$

By an obvious expansion, this is $\pm\theta(x, x_1)$, and the theorem follows from Theorem 10.1.

THEOREM 11.2. *If $\eta_1, \cdots, \eta_n$ are solutions of the Jacobi equations (2.24) vanishing at x_1, and the determinant*

$$d(x, x_1) = |\, \eta_j^i(x)\,|$$

is not identically zero, then its zeros $x_2 \neq x_1$ determine points on g conjugate to $\gamma(x_1)$.

The determinant $D(t, t_1)$ of Theorem 10.2 takes the form (if we use (11.2))

$$\begin{vmatrix} 0 & 1 \\ \eta_j^i(x) & \dot{\gamma}^i(x) \end{vmatrix}$$

which is clearly $\pm d(x, x_1)$. The theorem follows from Theorem 10.2.

We now return to the problem in parametric form.

12. There are three kinds of invariance which it is desirable (though not at all vital) to establish. We shall show that changing the vector function $p(t)$ does not affect conjugacy, and that the family of p-normal solutions is invariant under change of parameter and of coördinates.

LEMMA 12.1. *Let $p_i(t)$ and $\pi_i(t)$ be two systems of vectors both satisfying (5.1) and (5.2). If $\gamma(t_3)$ and $\gamma(t_4)$ are conjugate when p-normal solutions are used in the definition of conjugacy, they are conjugate when π-normal solutions are used. If there is a set of q linearly independent p-normal solutions vanishing at t_3 and t_4, then there is a set of q linearly independent π-normal solutions vanishing at t_3 and at t_4.*

We consider the second statement first. Let $\eta_1(t), \cdots, \eta_q(t)$ be linearly independent p-normal solutions of the Jacobi equations such that

$$\eta_k^i(t_3) = \eta_k^i(t_4) = 0 \qquad (i = 1, \cdots, n;\ k = 1, \cdots, q). \tag{12.1}$$

By Lemma 9.1, there are functions $\rho_k(t)$ $(k = 1, \cdots, q)$ such that the functions $\eta_k^i(t) - \rho_k(t)\dot{\gamma}^i(t)$ are π-normal solutions of the Jacobi equations which vanish at t_3 and at t_4. If they are not linearly independent, there exist numbers $c_1, \cdots, c_q$ such that

$$c_k[\eta_k^i(t) - \rho_k(t)\dot{\gamma}^i(t)] \equiv 0. \tag{12.2}$$

Let these be multiplied by the $p_i(t)$ and added. Since the $\eta_k(t)$ are p-normal, the first term vanishes, and by (5.2) we have $c_k\rho_k(t) \equiv 0$. Substituting this in (12.2) yields

$$c_k\eta_k^i(t) \equiv 0, \tag{12.3}$$

and this contradicts the linear independence of the $\eta_k(t)$.

In the special case $q = 1$, the result just established takes the form that if there is a non-identically vanishing p-normal solution vanishing at t_3 and t_4, then there is a non-identically vanishing π-normal solution vanishing at t_3 and t_4. This establishes the first statement in the lemma.

By Jacobi's theorem, if $y = y(t, e)$ is a family of extremals containing $\gamma(t)$ for $e = 0$, then $\eta(t) = y_e(t, 0)$ satisfies the Jacobi equations. For our purposes we need to know that if properly represented, this family yields a p-normal solution.

LEMMA 12.2. *Let $y = \bar{y}(\tau, e)$ be a family of curves [extremals] defined for τ on an interval $[T_1, T_2]$ containing $[t_1, t_2]$ in its interior and for all e near zero, and such that $\bar{y}(\tau, 0) = \gamma(\tau)$ $(t_1 \leqq \tau \leqq t_2)$. Let $\bar{y}(\tau, e)$ and $\bar{y}_\tau(\tau, e)$ be of class C^2. It is then possible to introduce a new parameter t by a function $\tau = \tau(t, e)$ in such a manner that*

(i) *the functions $\tau(t, e)$ are of class C^2 for t in $[t_1, t_2]$ and e near zero, and*

$$\tau(t, 0) \equiv t; \tag{12.4}$$

(ii) *the functions*

$$y(t, e) \equiv \bar{y}(\tau(t, e), e) \qquad (t_1 \leqq t \leqq t_2) \tag{12.5}$$

are of class C^2 for e near zero, and the derivatives $y_i(t, 0)$ are of class C^2;

(iii) *the derivatives $y_e(t, 0)$ are p-normal [solutions of the Jacobi equations].*

We prove the theorem omitting the bracketed words; the theorem including the bracketed words is an immediate consequence, by Jacobi's theorem. As before, we assume

$$p_i(t)\dot{\gamma}^i(t) \equiv 1,$$

and we define

$$\tau(t, e) = t - ep_i(t)\bar{y}_e^i(t, 0) \qquad (t_1 \leqq t \leqq t_2). \tag{12.6}$$

This is of class C^2, and τ is in $[T_1, T_2]$ if e is near zero, so for such e the function $y(t, e)$ is of class C^2. Equation (12.4) clearly holds, so

$$y_i(t, 0) = \bar{y}_\tau(t, 0),$$

which is of class C^2. Thus conclusions (i) and (ii) hold.

Recalling that $\bar{y}(t, 0) = \gamma(t)$ and using (12.4), (12.5) and (12.6), we obtain

$$\begin{aligned} y_e^i(t, 0) &= \bar{y}_\tau^i(t, 0)\tau_e(t, 0) + \bar{y}_e^i(t, 0) \\ &= -\dot{\gamma}^i(t)p_j(t)\bar{y}_e^j(t, 0) + \bar{y}_e^i(t, 0). \end{aligned}$$

Multiplying by $p_i(t)$ and summing on i yield

$$p_i(t)y_e^i(t, 0) = 0.$$

This completes the proof.

An immediate corollary is

LEMMA 12.3. *If in Lemma* 12.2 *the equations*

$$p_i(t_1)\bar{y}_e^i(t_1, 0) = p_i(t_2)\bar{y}_e^i(t_2, 0) = 0 \tag{12.7}$$

hold, then $\tau(t_1, e) \equiv t_1$ *and* $\tau(t_2, e) \equiv t_2$. *In this case we can take* $T_1 = t_1$, $T_2 = t_2$.

This follows at once from (12.6).

It is clear that we can regard e as a multipartite parameter $(e_1, \cdots, e_k)$, the right member of (12.6) being replaced by

$$t - e_h p_i(t)\bar{y}_{e_h}^i(t, 0),$$

summed for $h = 1, \cdots, k$. Lemmas 12.2 and 12.3 are still valid.

The invariance of the p-normal solutions under change of coördinates is an easy consequence of a theorem proved by Tucker.[14] However, we shall give an independent proof.

Let $\eta(t)$ be a p-normal solution of the Jacobi equations. Under a C^4 transformation of coördinates $z = z(y)$ the vector $\eta(t)$ transforms into

$$\zeta^i(t) = \frac{\partial z^i}{\partial y^j}\eta^j(t). \tag{12.8}$$

If $p(t)$ transforms into $q(t)$, from the invariance of $p_i\eta^i$ it follows that the equation

$$q_i(t)\zeta^i(t) = 0$$

holds. It remains to show that $\zeta(t)$ is a solution of the (transformed) Jacobi equations.

By known theorems on extremals, for every number e sufficiently near zero there is an extremal

$$y^i = y^i(t, e) \qquad (T_1 \leqq t \leqq T_2), \tag{12.9}$$

defined on an interval $[T_1, T_2]$ larger than $[t_1, t_2]$, such that the functions $y(t, e)$ and $y_t(t, e)$ are of class C^2, coincide with $\gamma(t)$ for $e = 0$, and satisfy the initial conditions

$$\begin{aligned} y^i(t_1, e) &= \gamma^i(t_1) + e\eta^i(t_1), \\ \dot{y}^i(t_1, e) &= \dot{\gamma}^i(t_1) + e\dot{\eta}^i(t_1). \end{aligned} \tag{12.10}$$

We suppose (as by Lemma 12.2 we may) that the parametric representation is such that

$$p_i(t)y_e^i(t, 0) \equiv 0. \tag{12.11}$$

[14] A. W. Tucker, *On tensor invariance in the calculus of variations*, Annals of Math., vol. 35(1934), pp. 341–350.

Then $y_e(t, 0)$ is a p-normal solution of the Jacobi equations and at t_1 coincides in value and derivative with $\eta(t)$. By Lemma 7.2,

$$y_e^i(t, 0) \equiv \eta^i(t). \tag{12.12}$$

After the transformation $z = z(y)$ let $F(y, r)$ assume the form $Q(z, s)$. Define

$$z^i(t, e) \equiv z^i(y(t, e)). \tag{12.13}$$

These functions are of class C^2. Moreover, along each of them the first variation of $\int Q\,dt$ with respect to any family of variations is zero, since the value of the integral remains unchanged by the transformation. Hence the curves (12.13) are extremals for the transformed integral. For $e = 0$ they contain the transform

$$z_0^i(t) \equiv z^i(\gamma(t)) \tag{12.14}$$

of our extremal g. Therefore by Jacobi's theorem the partial derivatives $z_e(t, 0)$ satisfy the Jacobi equations corresponding to $\int Q\,dt$. These derivatives are

$$z_e^i(t, 0) = \frac{\partial z^i}{\partial y^j}\, y_e^j(t, 0). \tag{12.15}$$

Comparing (12.15), (12.12) and (12.8), we see that $\zeta^i(t)$ is a solution of the Jacobi equations, and the proof is complete.

The invariance of p-normal solutions under change of parameter is easily established by computation. It can also be established as above. For in the family (12.9) let us introduce a new parameter τ by the equation

$$t = t(\tau) \qquad (\tau_1 \leqq \tau \leqq \tau_2),$$

where $t(\tau)$ is of class C^2 and $t'(\tau) > 0$. The family assumes the form

$$y^i = Y^i(\tau, e) = y^i(t(\tau), e) \qquad (\tau_1 \leqq \tau \leqq \tau_2). \tag{12.16}$$

These are still extremals, so their partial derivatives

$$Y_e^i(\tau, 0) = y_e^i(t(\tau), 0) \tag{12.17}$$

satisfy the Jacobi equations. But by (12.12)

$$y_e^i(t(\tau), 0) = \eta^i(t(\tau)). \tag{12.18}$$

Hence $\eta^i(t(\tau))$ satisfies the Jacobi equations.

By (5.3), after the change of parameter the functions $p_i(t)$ assume the form $q_i(\tau)$, where

$$q_i(\tau) = k(\tau)p_i(t(\tau)), \qquad k \neq 0.$$

Hence

$$q_i(\tau)\dot{Y}^i(\tau, 0) = k(\tau)p_i(t(\tau))\cdot\dot{\gamma}^i(t(\tau))\cdot t'(\tau) \neq 0,$$

and (5.2) continues to hold. By (5.3), $q(\tau)$ is of class C^2. Also

$$q_i(\tau)\eta^i(t(\tau)) = k(\tau)p_i(t(\tau))\eta^i(t(\tau)) = 0,$$

so the $\eta^i(t(\tau))$ are q-normal solutions of the transformed Jacobi equations.

Next we take up the proof of Morse's index theorem.

13. We continue to assume that g is a non-singular extremal with $(\gamma(t), \dot{\gamma}(t))$ interior to R. Furthermore, we shall assume that the Legendre condition holds. In the presence of non-singularity, this is equivalent to assuming the strengthened Legendre condition:

(III′) *$u^iF^0_{\tau^i\tau^j}u^j > 0$ for all t in $[t_1, t_2]$ and all u not of the form $u^i = k\dot{\gamma}^i(t)$.*

Let

$$t_1 = a_0 < a_1 < \cdots < a_p < a_{p+1} = t_2$$

be a sequence of values of t such that no two points in any one interval $[a_i, a_{i+1}]$ are conjugate. For $q = 1, \cdots, p$ let M_q be an $(n-1)$-dimensional manifold defined by functions

$$y^i = y^i_q(u^1, \cdots, u^{n-1}) \tag{13.1}$$

of class C^3, such that

$$y^i_q(0) = \gamma^i(a_q). \tag{13.2}$$

We assume that each matrix

$$\left(\frac{\partial y^i_q}{\partial u^j}\right) \tag{13.3}$$

has rank $n-1$ at $u = 0$, and furthermore we assume that M_q is not tangent to the extremal g at a_q. That is, the vector $\dot{\gamma}(a_q)$ is not a linear combination of the columns of (13.3). Given a sequence of p $(n-1)$-tuples of numbers, we write them consecutively in the form

$$(z^1, \cdots, z^{p(n-1)}) = (u^1_1, \cdots, u^{n-1}_1, \cdots, u^1_p, \cdots, u^{n-1}_p). \tag{13.4}$$

We may thus consider a set z as defining (if near enough to 0) a point

$$P_1 = y_1(u^1_1, \cdots, u^{n-1}_1)$$

on M_1, a point P_2 on M_2, etc.

Since $\gamma(a_{i+1})$ is not conjugate to $\gamma(a_i)$ $(i = 0, 1, \cdots, p)$, we know that for z near enough to 0 the points $\gamma(t_1)$, P_1, P_2, $\cdots$, P_p, $\gamma(t_2)$ can be joined in that order by arcs of extremals.

For each q let p_q be a vector orthogonal to M_q, that is, orthogonal to the columns of (13.3). Since $\dot{\gamma}(a_q)$ is not a linear combination of these columns, we have

$$p_{i,q}\dot{\gamma}^i(a_q) \neq 0,$$

not summed on q. By reversing p_q if necessary we can obtain

$$p_{i,q}\dot{\gamma}^i(a_q) > 0 \qquad (q = 1, \cdots, p). \tag{13.5}$$

Also we choose vectors p_0, p_{p+1} such that

$$p_{i,0}\dot{\gamma}^i(a_0) > 0, \qquad p_{i,p+1}\dot{\gamma}^i(a_{p+1}) > 0. \tag{13.6}$$

It is now easy to determine a function $p(t)$ of class C^2 such that

$$\begin{aligned} &p_i(t)\dot{\gamma}^i(t) > 0 && (t_1 \leqq t \leqq t_2), \\ &p_i(a_q) = p_{i,q} && (q = 0, 1, \cdots, p+1). \end{aligned} \tag{13.7}$$

For each z, let $E(z)$ be the broken extremal with vertices at the points $\gamma(t_1)$, $P_1, \cdots, P_p, \gamma(t_2)$ determined by z. Let us restrict our attention to a particular interval $[a_k, a_{k+1}]$. There is no loss of generality in assuming that the arc of $E(z)$ between P_k and P_{k+1} is represented by a function

$$y^i = y^i(t, z) \qquad (a_k \leqq t \leqq a_{k+1}) \tag{13.8}$$

such that $y(a_k, z)$ is P_k and $y(a_{k+1}, z)$ is P_{k+1}; that is,

$$\begin{aligned} y^i(a_k, z) &= y^i_k(u^1_k, \cdots, u^{n-1}_k), \\ y^i(a_{k+1}, z) &= y^i_{k+1}(u^1_{k+1}, \cdots, u^{n-1}_{k+1}). \end{aligned} \tag{13.9}$$

Consequently, each of the derivatives

$$\frac{\partial y^i(a_k, z)}{\partial z^h}, \qquad \frac{\partial y^i(a_{k+1}, z)}{\partial z^h}$$

either is one of the derivatives

$$\frac{\partial y^i_k}{\partial u^j_k}, \qquad \frac{\partial y^i_{k+1}}{\partial u^j_{k+1}} \qquad (k \text{ not summed})$$

or else is 0, according as z^h is one of the u^i_k or u^i_{k+1} or is not one of them. In any case, by hypothesis we have

$$\begin{aligned} p_i(a_k)\,\frac{\partial y^i(a_k, z)}{\partial z^h} &= 0, \\ p_i(a_{k+1})\,\frac{\partial y^i(a_{k+1}, z)}{\partial z^h} &= 0, \end{aligned} \tag{13.10}$$

not summed on k. Thus Lemmas 12.3 and 12.2 can be applied, and we may suppose that the parameter t, on the interval $[a_k, a_{k+1}]$, is already the parameter whose existence is established in those lemmas. This yields

LEMMA 13.1. *The broken extremals $E(z)$ can be represented by functions $y = y(t, z)$ $(t_1 \leqq t \leqq t_2)$ such that $y(a_q, z) = P_q$ $(q = 1, \cdots, p)$. On each interval $[a_k, a_{k+1}]$ these functions are of class C^2. The functions*

$$a^i_\alpha(t) \equiv \left.\frac{\partial y^i(t, z)}{\partial z^\alpha}\right|_{z=0} \tag{13.11}$$

define continuous curves with corners at the values a_q of t and of class C^2 between corners. Between corners, the functions (13.11) *are p-normal solutions of the Jacobi equations.*

With the notation (13.11), the statements that matrix (13.3) has rank $n-1$ at $u=0$ and that $\dot\gamma(a_q)$ is not a linear combination of its columns can be written in the form

$$|\, a^i_\alpha(a_q)\,\dot\gamma^i(a_q)\,| \neq 0 \qquad (\alpha = (n-1)(q-1)+1, \cdots, (n-1)q;\ i = 1, \cdots, n). \tag{13.12}$$

14. In studying the index theorem we shall for brevity refer in large part to the proof given by Morse,[15] indicating changes required to bring them in the scope of our methods. We replace all references to the RJE (restricted Jacobi equations) of Morse by references to our equations (7.3).

THEOREM[16] 14.1. *If $\Theta(t)$ is a determinant whose columns are n linearly independent solutions of* (7.3) *which vanish at $t=t_1$, the order of vanishing of $\Theta(t)$ at any point $t=a$ equals the nullity ν of $\Theta(a)$.*

With Morse, we define the order of a as a conjugate point of t_1 to be the order of vanishing of $\Theta(t)$ at $t=a$.

The proof of Morse needs no change if we assume, as we may without loss of generality, that $p_i\dot\gamma^i \equiv 1$. (This is needed in equation (2.19) of Morse.)

COROLLARY. *The maximum number of linearly independent p-normal solutions which vanish at t_1 and at a conjugate point a equals the order of a as a conjugate point of t_1.*

This is Corollary 2.2 of Morse, except that we have "p-normal solutions" instead of "solutions of (7.3)". Let $\eta_1, \cdots, \eta_\nu$ be solutions of (7.3) vanishing at t_0 and at a. By Lemma 7.1 the functions

$$p_i(t)\eta^i_k(t) \qquad (k = 1, \cdots, \nu)$$

are linear in t. They vanish at t_1 and at $a \neq t_1$, so they vanish identically, and the η_k are p-normal.

It follows that the maximum order of the conjugate points is $n-1$, since there are at most $n-1$ p-normal solutions in an independent set vanishing at t_1.

Let $J(z)$ be the value of the integral $J(C)$ along the broken extremal $E(z)$ determined by z. Then we have

THEOREM 14.2 (Index theorem). *The point $z=(0, \cdots, 0)$ is a critical point of $J(z)$ with an index equal to the number of conjugate points of $t=t_1$ on g preceding $t=t_2$, and a nullity equal to the order of $t=t_2$ as a conjugate point of $t=t_1$ on g.*

[15] See footnote 4.

[16] Morse, loc. cit. (footnote 4), Theorem 22.

On the extremals $E(z)$ we introduce the parameter of Lemma 13.1. Beginning with Morse's equation (3.5), we proceed to the end of §3. The first paragraph of §4 is supplanted by our Lemma 13.1.

In place of Morse's Lemma 4.2 we have

LEMMA 14.3. *A necessary and sufficient condition that a set $z \neq 0$ be a critical point of $Q(z)$ is that the function*

$$\eta^i(t) = z^a a_a^i(t)$$

(cf. (13.11)*) determined by z be a p-normal solution.*

The sufficiency is easy to establish. If $\eta(t)$ is a p-normal solution, it is of class C^2, so that

$$\Omega_{p^i}(\eta(a_q + 0), \dot{\eta}(a_q + 0)) = \Omega_{p^i}(\eta(a_q - 0), \dot{\eta}(a_q - 0)).$$

Hence Morse's equation (4.6)′ holds, and z is critical. Conversely, Morse shows in the proof of his Lemma 4.2 that if z is critical, the equations

$$\dot{\eta}^i(a_q + 0) - \dot{\eta}^i(a_q - 0) = c\dot{\gamma}^i(a_q) \tag{14.1}$$

hold for some c. By our Lemma 13.1 the function $\eta(t)$ is p-normal, so that $p_i\eta^i \equiv 0$. Differentiation yields

$$\dot{p}_i\eta^i(a_q) + p_i\dot{\eta}^i(a_q + 0) = \dot{p}_i\eta^i(a_q) + p_i\dot{\eta}^i(a_q - 0) = 0,$$

whence

$$p_i(a_q)[\dot{\eta}^i(a_q + 0) - \dot{\eta}^i(a_q - 0)] = 0. \tag{14.2}$$

If we multiply by $p_i(a_q)$ in (14.1) and sum, we find from (14.2) and (5.2) that $c = 0$, so that η has no corner at a_q.

Thus η is of class C^1. Since it satisfies (7.3) on each interval $[a_q, a_{q+1}]$, it has a continuous second derivative also and satisfies (7.3) on all of $[t_1, t_2]$. Since it is p-normal by Lemma 13.1, the lemma is established.

This lemma simplifies the proof of Morse's Theorem 4.1, which is our

THEOREM 14.4. *The nullity of $Q(z)$ equals the order of $t = t_2$ as a conjugate point of $t = t_1$.*

The nullity of the form $Q(z)$ is the nullity (order minus rank) of its matrix. If we temporarily write

$$Q(z) = b_{ij}z^i z^j \qquad (i, j = 1, \cdots, p(n-1)),$$

then z is a critical point for Q if and only if

$$b_{ij}z^j = 0 \qquad (i = 1, \cdots, p(n-1)).$$

The maximal number of linearly independent solutions of such a system of homogeneous equations is equal to the nullity of the matrix of coefficients, so

(A) *the nullity ν of Q equals the maximal number of linearly independent sets $z_1, \cdots, z_\nu$ at each of which Q is critical.*

By our Lemma 14.3 and Lemma 4.1 of Morse,

(B) *the maximal number ν of linearly independent critical sets $z_1, \cdots, z_\nu$ equals the maximal number of linearly independent p-normal solutions of the form $z^\alpha a_\alpha^i(t)$.*

Next we shall establish the following statement.

(C) *All p-normal solutions $\eta(t)$ which vanish at t_1 and at t_2 are of the form $z^\alpha a_\alpha^i(t)$.*

Since the determinant (13.12) is not zero, there are numbers z^α ($\alpha = (n-1)(q-1)+1, \cdots, (n-1)q$) and d_q such that

$$\eta^i(a_q) = c^\alpha a_\alpha^i(a_q) + d_q \gamma^i(a_q). \tag{14.3}$$

If we multiply by $p_i(a_q)$ and sum on i, the left member vanishes because $\eta(t)$ is p-normal, and the first term on the right vanishes by (13.10) and (13.11). Therefore we have $d_q = 0$.

All the numbers $a_\alpha^i(a_q)$ are zero except those with $\alpha = (n-1)(q-1)+1, \cdots, (n-1)q$, so by (14.3) the function

$$\eta^i(t) - c^\alpha a_\alpha^i(t) \qquad (\alpha = 1, \cdots, (n-1)p)$$

vanishes at each a_q. It is a p-normal solution on each interval (a_q, a_{q+1}), so it cannot be different from zero between these values; if it were, they would be conjugate, contrary to hypothesis. Therefore $\eta^i(t) \equiv c^\alpha a_\alpha^i(t)$, as was to be proved.

Comparing statements (A), (B) and (C) with the corollary to Theorem 14.1, our present theorem is established.

Now we follow the proof of Morse from Theorem 4.2 to the end of the paper without change, at the end of which process the proof of the index theorem (our Theorem 14.2) is complete.

15. The applications of p-normal variations discussed above are not exhaustive. They can equally well be used in connection with natural isoperimetric conditions,[17] and also in studying problems with variable end points. But it is not desirable to lengthen this paper further by going into greater detail.

A possible alteration of the requirements on our functions $p(t)$ is to replace (5.3) by the condition that under change of parameter $t = t(\tau)$, the functions $p_i(t)$ are replaced by functions $q_i(\tau)$ for which (5.2) remains valid. With this weakening of (5.3) it is not necessarily true that the class of p-normal variations is invariant under change of parameter, but as in proving Lemma 12.1 we can show that the location and order of conjugate points remains unchanged. This proof is of course not necessary, in view of the corollary to Theorem 10.2 (for location of conjugate points) and Theorem 14.4 (for location and order).

UNIVERSITY OF VIRGINIA.

[17] Birkhoff and Hestenes, loc. cit. (footnote 3).

PERIODICALS PUBLISHED BY DUKE UNIVERSITY

American Literature. A quarterly journal devoted to research in American Literature, published with the coöperation of the American Literature Group of the Modern Language Association of America. Subscription, $4.00 per year. Back volumes, $5.00 each.

Character and Personality. A psychological quarterly devoted to studies of behavior and personality. Subscription, $2.00 per year. The first number was published September, 1932.

Contributions to Psychological Theory. A monograph series dealing with problems of psychological theory in the widest sense, including their relations to other fields of inquiry. The monographs appear irregularly. Subscription, $5.00 per volume of approximately 450 pages.

Duke Mathematical Journal.

Ecological Monographs. A quarterly journal devoted to the publication of original researches of ecological interest from the entire field of biological science. Subscription, $6.00 per year. The first number was published January, 1931.

Hispanic American Historical Review. A quarterly review dealing with the history of the Latin-American countries. Subscription, $4.00 per year.

The Journal of Parapsychology. A scientific quarterly dealing with telepathy, clairvoyance, and other parapsychological problems. Subscription, $3.00 per year.

Law and Contemporary Problems. A quarterly published by the School of Law, presenting in each issue a symposium on a problem of current importance having significant legal aspects. Subscription, $2.00 per year. The first number was published September, 1933.

The South Atlantic Quarterly. A magazine of modern opinion and discussion, founded in 1902. Subscription, $3.00 per year.

The Southern Association Quarterly. As official organ of the Southern Association of Colleges and Secondary Schools, it contains the proceedings of the annual meeting, together with much additional material directly related to the work of the Association. Subscription, $4.00 per year.

DUKE UNIVERSITY PRESS
DURHAM, NORTH CAROLINA

CONTENTS

The ergodic theorem. By Norbert Wiener 1
The mean ergodic theorem. By Garrett Birkhoff 19
The analytic prolongation of a minimal surface over a rectilinear segment of its boundary. By Jesse Douglas 21
Surfaces in four-space of constant curvature. By Nathaniel Coburn ... 35
The characterization of flat projective spaces by their differential invariants, and certain related problems. By Robert R. Singleton 57
The Waring problem with summands x^m, $m \geqq n$. By Mary Haberzetle. 69
Isometric embedding of flat manifolds in Euclidian space. By C. Tompkins 89
Simultaneous invariants of a complex and subcomplex. By C. E. Clark 97
Lineal element transformations of space for which normal congruences of curves are converted into normal congruences. By Edward Kasner. 72
The subgroup of order n of a transitive group of degree n and class $n - 1$. By Louis Weisner 86
Identities analogous to Ramanujan's identities involving the partition function. By Herbert S. Zuckerman 95
Hereditary arc sums. By O. G. Harrold, Jr 111
A generalization of Picard's and related theorems. By Raphael M. Robinson 125
Certain equivalences and subsets of a plane. By F. Burton Jones 133
The maps determined by the principal curves associated with five and six points in the plane. By C. Ronald Cassity 145
Functions which assume rational values at rational points. By J. W. Green 154
A remark on the normal decompositions of groups. By Oystein Ore .. 172
On imbedding a space in a complete space. By L. W. Cohen 174
The Jacobi condition and the index theorem in the calculus of variations. By E. J. McShane 184